

Dynamic Portfolio Choice with Linear Rebalancing Rules*

Ciamac C. Moallemi

Graduate School of Business

Columbia University

email: ciamac@gsb.columbia.edu

Mehmet Sağlam

Bendheim Center for Finance

Princeton University

email: msaglam@princeton.edu

Current Revision: March 25, 2015

Abstract

We consider a broad class of dynamic portfolio optimization problems that allow for complex models of return predictability, transaction costs, trading constraints, and risk considerations. Determining an optimal policy in this general setting is almost always intractable. We propose a class of linear rebalancing rules and describe an efficient computational procedure to optimize with this class. We illustrate this method in the context of portfolio execution and show that it achieves near optimal performance. We consider another numerical example involving dynamic trading with mean-variance preferences and demonstrate that our method can result in economically large benefits.

*Sağlam acknowledges support from the Eugene M. Lang Doctoral Student Grant. Moallemi acknowledges the support of NSF grant CMMI-1235023. We are grateful for helpful comments from David Brown, Sylvain Champonnois (discussant), Michael Sotiropoulos, and conference participants at the 6th Annual Conference on Advances in the Analysis of Hedge Fund Strategies at Imperial College London.

1. Introduction

Dynamic portfolio optimization has been a central and essential objective for institutional investors in active asset management. Real world portfolio allocation problems of *practical* interest have a number of common features:

Return predictability. At the heart of active portfolio management is the fact that a manager will seek to predict future asset returns. Such predictions are not limited to simple unconditional estimates of expected future returns, but often involve predictions on short- and long-term expected returns using complex models based on observable return predicting factors.

Transaction costs. Trading costs in dynamic portfolio management can arise from sources ranging from the bid-offer spread or execution commissions to price impact, where the manager's own trading affects the subsequent evolution of prices.

Portfolio or trade constraints. Often times managers cannot make arbitrary investment decisions, but rather face exogenous constraints on their trades or their resulting portfolio. Examples of this include short-sale constraints, leverage constraints, or restrictions requiring market neutrality (or specific industry neutrality).

Risk aversion. Portfolio managers seek to control the risk of their portfolios. In practical settings, risk aversion is not accomplished by the specification of an abstract utility function. Rather, managers specify limits or penalties for multiple summary statistics that capture aspects of portfolio risk which are easy to interpret and are known to be important. For example, a manager may both be interested in the risk of the portfolio value changing over various intervals of time, including for example, both short intervals (e.g., daily or weekly risk), as well as risk associated with the terminal value of the portfolio. Such single-period risk can be measured a number of ways (e.g., variance, value-at-risk). A manager might further be interested in multi-period measures of portfolio risk, for example, the maximum drawdown of the portfolio.

Significantly complicating the analysis of portfolio choice is that the underlying problem is *multi-period*. Here, in general, the decision made by a manager at a given instant of time might depend on all information realized up to that point. Traditional approaches to multi-period portfolio choice, dating back at least to the work of Merton (1971), have focused on analytically determining the optimal dynamic policy. While this work has brought forth important structural insights, it is fundamentally quite restrictive: exact analytical solutions require very specific assumptions about investor objectives and market dynamics. These assumptions cannot accommodate flexibility in, for example, the return generating process, trading frictions, and constraints, and are often practically unrealistic. Absent such

restrictive assumptions, analytical solutions are not possible. Motivated by this, much of the subsequent academic literature on portfolio choice seeks to develop modeling assumptions that allow for analytical solutions, however the resulting formulations are often not representative of real world problems of practical interest. Further, because of the ‘curse-of-dimensionality’, exact numerical solutions are often intractable in cases of practical interest, where the universe of tradeable assets is large.

In search of tractable alternatives, many practitioners eschew multi-period formulations. Instead, they consider portfolio choice problems in a myopic, single-period setting, when the underlying application is clearly multi-period (e.g., Grinold and Kahn, 1999). Another tractable possibility is to consider portfolio choice problems that are multi-period, but without the possibility of recourse. Here, a fixed set of deterministic decisions for the entire time horizon is made at the initial time. Both single-period and deterministic portfolio choice formulations are quite flexible and can accommodate many of the features described above. They are typically applied in a quasi-dynamic fashion through the method of *model predictive control*. Here, at each time period, the simplified portfolio choice problem is re-solved based on the latest available information.

While these simplified approaches are extremely flexible and have been broadly adopted in practice, these methods have important flaws. In general, such methods are heuristics; in order to achieve tractability, they neglect the explicit consideration of the possibility of future recourse. Hence, these methods may be significantly sub-optimal. Moreover, single-period formulations, which are the most popular among practitioners, pose a number of additional challenges. In general, they do not effectively manage transaction costs; re-solving a single-period model repeatedly causes portfolio churn. They are also difficult to apply in situations where returns are predicted across multiple time horizons. Ideally, an investor should be very responsive to short-term predictions that will be realized quickly, while responding less aggressively to long-term predictions where there is time to work into a position. It is not clear how to accommodate this in a single-period setting that allows only a single choice of time horizon. In general, practitioners adopt *ad hoc* heuristics to address these issues. For example, one can introduce artificial transaction costs to limit portfolio churn, or one can artificially scale return predictors based on their relative horizons.

Another tractable alternative is the formulation of portfolio choice problems as *linear quadratic control* (e.g., Hora, 2006; Gârleanu and Pedersen, 2013). Since the 1950’s, linear quadratic control problems have been an important class of tractable multi-period optimal control problems. In the setting of portfolio choice, if the return dynamics are linear, transaction costs and risk aversion penalties can be decomposed into per-period quadratic functions, and security holdings and trading decision are unconstrained, then these methods

apply. However, there are many important problem cases that simply do not fall into the linear quadratic framework.

In this paper, our central innovation is to propose a framework for multi-period portfolio optimization, which admits a broad class of problems including many features described earlier. Our formulation maintains tractability by restricting the problem to determining the best policy out of a restricted class of *linear rebalancing policies*. Such policies allow planning for future recourse, but only of a form that can be parsimoniously parameterized in a specific affine fashion. In particular, the contributions of this paper are as follows:

First, we define a flexible, general setting for portfolio optimization. Our setting allows for very general dynamics of asset prices, with arbitrary dependence on the history of ‘return-predictive factors’. We allow for any convex constraints on trades and positions. Finally, the objective is allowed to be an arbitrary concave function of the sample path of positions. Our framework admits, for example, many complex models for transaction costs or risk aversion. We can consider both traditional problem formulations for portfolio optimization (e.g., maximization of expected terminal utility of wealth) as well as formulations more popular with practitioners (e.g., maximization of expected wealth subject to risk constraints).

Second, our portfolio optimization problem is computationally tractable. In our setting, determining the optimal linear rebalancing policy is a convex program. Convexity guarantees that the globally optimal policy can be tractably found in general. This is in contrast to non-convex portfolio choice parameterizations (e.g., Brandt et al., 2009), where only local optimality can be guaranteed.

In our case, numerical solutions can be obtained via, for example, sample average approximation or stochastic approximation methods (see, e.g., Shapiro, 2003; Nemirovski et al., 2009). These methods can be applied in a data-driven fashion, with access only to simulated trajectories and without an explicit model of system dynamics. In a number of instances where the factor and return dynamics are driven by Gaussian uncertainty, we illustrate that our portfolio optimization problem can be reduced to a standard form of convex optimization program, which can be solved with off-the-shelf commercial optimization solvers.

Third, our class of linear rebalancing policies subsumes many common heuristic portfolio policies. Both single-period and deterministic policies are special cases of linear rebalancing policies, however linear rebalancing policies are a broader class. Hence, the optimal linear rebalancing policy will outperform policies from these more restricted classes. Further, our method can also be applied in the context of model predictive control. Also, portfolio optimization problems that can be formulated as linear quadratic control also fit in our setting, and their optimal policies are linear rebalancing rules.

Finally, we demonstrate the practical benefits of our method in two examples: optimal ex-

ecution with trading constraints and dynamic trading with mean-variance preferences. First, we consider an optimal execution problem where an investor seeks to liquidate a position over a fixed time horizon, in the presence of transaction costs and a model for predicting returns. We further introduce linear inequality constraints that require the trading decisions to only be sales; such sale-only constraints are common in agency algorithmic trading. The resulting optimal execution problem does not admit an exact solution. Hence, we compare the best linear policy to a number of tractable alternative approximate policies, including a deterministic policy, model predictive control, and a projected variation of the linear quadratic control formulation of Gârleanu and Pedersen (2013). We demonstrate that the best linear policy achieves superior performance to the alternatives. Moreover, we compute a number of upper bounds on the performance of *any* policy in the problem at hand. Using these upper bounds, we see that the best linear policy is near optimal, with a gap of at most 5%. Our sensitivity analysis shows that the percentage improvement obtained using linear rebalancing rules can be up to 18% when compared with the best alternative policy. Second, we consider a dynamic trading problem where an investor with mean-variance preferences makes intraday trading decisions in the presence of return predictability. Using the same model calibration in the optimal execution example, we illustrate that the gains from using our best linear policy can be economically substantial when the model does not fall within realm of linear-quadratic formulation. Moreover, our sensitivity analysis reveals that this outperformance is robust to different model calibrations and can provide an improvement of 72% when benchmarked against a trading rule based on a linear quadratic formulation.

Literature review. Our paper is related to two different strands of literature: the literature of dynamic portfolio choice with return predictability and transaction costs, and the literature on the use of linear decision rules in the optimal control problems.

First, we consider the literature on dynamic portfolio choice. This vast body of work begins with the seminal paper of Merton (1971). Following this paper, there has been a significant literature aiming to incorporate the impact of various frictions, such as transaction costs, on the optimal portfolio choice ¹. Liu and Loewenstein (2002) study the optimal trading strategy for a constant relative risk aversion (CRRA) investor in the presence of transaction costs and obtain closed-form solutions when the finite horizon is uncertain. De-

¹The work of Constantinides (1986) is an early example that studies the impact of proportional transaction costs on the optimal investment decision and the liquidity premium in the context of the capital asset pricing model (CAPM). Davis and Norman (1990), Dumas and Luciano (1991), and Shreve and Soner (1994) provide the exact solution for the optimal investment and consumption decision by formally characterizing the trade and no-trade regions. One drawback of these papers is that the optimal solution is only computed in the case of a single stock and bond. For a survey on this literature, see Cvitanic (2001). Liu (2004) extends these results to multiple assets with fixed and proportional transaction costs in the case of uncorrelated asset prices.

temple et al. (2003) develop a simulation-based methodology for optimal portfolio choice in complete markets with complex state dynamics.

There is also a significant literature on portfolio optimization that incorporates return predictability (see, e.g., Campbell and Viceira, 2002). Balduzzi and Lynch (1999) and Lynch and Balduzzi (2000) illustrate the impact of return predictability and transaction costs on the utility costs and the optimal rebalancing rule by discretizing the state space of the dynamic program. With a similar state space discretization, Lynch and Tan (2010) model the dynamic portfolio decision with multiple risky assets under return predictability and transaction costs, and provide numerical experiments with two risky assets.

Much of the aforementioned literature seeks to find the best rebalancing policy out of the universe of all possible rebalancing policies. As discussed earlier, this leads to highly restrictive modeling primitives. On the other hand, our work is in the spirit of Brandt et al. (2009), who allow for broader modeling flexibility at the expense of considering a restricted class of rebalancing policies. They parameterize the rebalancing rule as a function of security characteristics and estimate the parameters of the rule from empirical data without modeling the distribution of the returns and the return predicting factors. Even though our approach is also a linear parameterizations of return predicting factors, there are fundamental differences between our approach and that of Brandt et al. (2009). First, the class of linear policies we consider is much larger than the specific linear functional form in Brandt et al. (2009). In our approach the parameters are time-varying and cross-sectionally different for each security. Second, the extensions provided in Brandt et al. (2009) for imposing positivity constraints and transaction costs are ad-hoc and cannot be generalized to arbitrary convex constraints or transaction cost functions. Finally, the objective function of Brandt et al. (2009) is a non-convex function of the policy parameters. Hence, it is not possible, in general to obtain the globally optimal set of parameters. Our setting, on the other hand, is convex, and hence globally optimal policies can be determined efficiently. Brandt and Santa-Clara (2006) use a different approximate policy for the optimal solution that invests in conditional portfolios, which invest in each asset an amount proportional to conditioning variables. Furthermore, Brandt et al. (2005) compute approximate portfolio weights using a Taylor expansion of the value function and approximating conditional expected returns as affine parameterizations of nonlinear functions.

Gârleanu and Pedersen (2013) achieve a closed-form solution for a model with linear dynamics for return predictors, quadratic functions for transaction costs, and quadratic penalty terms for risk². However, the analytic solution is highly sensitive to the quadratic

²Boyd et al. (2012) consider an alternative generalization of the linear-quadratic case, using ideas from approximate dynamic programming. Glasserman and Xu (2011) develop a linear-quadratic formulation for

cost structure with linear dynamics (see, e.g., Bertsekas, 2000). This special case cannot handle any inequality constraints on portfolio positions, non-quadratic transactions costs, or more complicated risk considerations. On the other hand, our approach can be implemented efficiently in these realistic scenarios and provides more flexibility in the objective function of the investor and the constraints that the investor faces.

Second, there is also a literature on the use of linear decision rules in optimal control problems. This approximation technique has attracted considerable interest recently in robust and two-stage adaptive optimization context ³. In this strand of literature, we believe the closest works to the methodology described in our paper are Calafiore (2009) and Skaf and Boyd (2010). Both of these papers use linear decision rules to address dynamic portfolio choice problems with proportional transaction costs without return predictability. Calafiore (2009) computes lower and upper bounds on the expected transaction costs and solves two convex optimization problems to get upper and lower bounds on the optimal value of the simplified dynamic optimization program with linear decision rules. On the other hand, Skaf and Boyd (2010) study the dynamic portfolio choice problem as an application to their general methodology of using affine controllers on convex stochastic programs. They first linearize the dynamics of the wealth process and then solve the resulting convex optimization via sampling techniques. The foremost difference between our approach and these papers is the modeling of return predictability. Hence, the optimal rebalancing rule in our model is a linear function of the predicting factors. Furthermore, we derive exact reductions to deterministic convex programs in the cases of proportional and nonlinear transaction costs.

2. Dynamic Portfolio Choice with Return Predictability and Transaction Costs

We consider a dynamic portfolio choice problem allowing general models for the predictability of security returns and for trading frictions. The number of investable securities is N , time is discrete and indexed by $t = 1, \dots, T$, where T is the investment horizon. Each security i has a price change of $r_{i,t+1}$ from time t to $t + 1$.

We collect these price changes in the return vector $r_{t+1} \triangleq (r_{1,t+1}, \dots, r_{N,t+1})$. We assume that the investor has a predictive model of future security returns, and that these predictions are made through a set of K *return-predictive factors*. These factors could be security-specific

portfolio optimization that offers robustness to modeling errors or mis-specifications.

³(See, e.g., Ben-Tal et al., 2004, 2005; Chen et al., 2007, 2008; Bertsimas et al., 2010; Bertsimas and Goyal, 2011). Shapiro and Nemirovski (2005) illustrate that linear decision rules can reduce the complexity of multistage stochastic programming problems. Kuhn et al. (2009) proposes an efficient method to estimate the loss of optimality incurred by linear decision rule approximation.

characteristics such as the market capitalization of the stock, the book-to-market ratio of the stock, the lagged twelve month return of the stock (see, e.g., Fama and French, 1996; Goetzmann and Jorion, 1993). Alternatively, they could be macroeconomic signals that affect the return of each security, such as inflation, treasury bill rate, industrial production (see, e.g., Chen et al., 1986). We denote by $f_t \in \mathbb{R}^K$ the vector of factor values at time t . Under the following assumption, we allow for very general dynamics, possibly nonlinear and with a general dependence on history, for the evolution of returns and factors:

Assumption 1 (General return and factor dynamics). *Over a complete filtered probability space given by $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbf{P})$, we assume that factors and returns evolve according to*

$$f_{t+1} = G_{t+1}(f_t, \dots, f_1, \epsilon_{t+1}), \quad r_{t+1} = H_{t+1}(f_t, \epsilon_{t+1}),$$

for each time t . Here, $G_{t+1}(\cdot)$ and $H_{t+1}(\cdot)$ are known functions that describe the evolution of the factors and returns in terms of the history of factor values and the exogenous i.i.d. disturbances ϵ_{t+1} . We assume that the filtration $\mathbb{F} \triangleq \{\mathcal{F}_t\}_{t \geq 0}$ is the natural filtration generated by the exogenous noise terms $\{\epsilon_t\}$.

Note that we choose to describe the evolution of asset prices in our framework in terms of absolute price changes, and we will also refer to these as (absolute) returns. This choice is purely notational and is without loss of generality: since the return dynamics specified by Assumption 1 allow for an arbitrary dependence on history, our framework also admits, for example, models which describe the percentage return of each security. Example 1 in Section 2.1 illustrates such a model.

Let $x_{i,t}$ denote the number of shares that the investor holds in the i th security over the time period t . We collect the portfolio holdings across all securities at time t in the vector $x_t \triangleq (x_{1,t}, \dots, x_{N,t})$, and we denote the fixed initial portfolio of the investor by x_0 . Similarly, let the trade vector $u_t \triangleq (u_{1,t}, \dots, u_{N,t})$ denote the amount of shares that the investor wants to trade at the beginning of the t th period, when he inherits the portfolio x_{t-1} from the prior period and observes the latest realization of factor values f_t . Consequently, we have the following linear dynamics for our position and trade vector: $x_t = x_{t-1} + u_t$, for each t .

Let the entire sample path of portfolio positions, factor realizations, and security returns be denoted by $\mathbf{x} \triangleq (x_1, \dots, x_T)$, $\mathbf{f} \triangleq (f_1, \dots, f_T)$, and $\mathbf{r} \triangleq (r_2, \dots, r_{T+1})$, respectively. Similarly, the sample path of trades over time is denoted by $\mathbf{u} = (u_1, \dots, u_T)$. We make the following assumption on feasible sample paths of trades:

Assumption 2 (Convex trading constraints). *The sample path of trades \mathbf{u} are restricted to the non-empty, closed, and convex set $U \subseteq \mathbb{R}^N \times \dots \times \mathbb{R}^N$.*

The investor's trading decisions are determined by a policy π that selects a sample path of trades \mathbf{u} in U for each realization of \mathbf{r} and \mathbf{f} . We let \mathcal{U} be the set of all policies. We assume that the investor's trading decisions are non-anticipating in that the trade vector u_t in period t depends only on what is known at the beginning of period t . Formally, we require policies to be adapted to the filtration \mathbb{F} , such that a policy's selection of the trade vector u_t at time t must be measurable with respect to \mathcal{F}_t . Let $\mathcal{U}_{\mathbb{F}}$ be the set of all non-anticipating policies.

The objective of the investor is to select a policy $\pi \in \mathcal{U}_{\mathbb{F}}$ that maximizes the expected value of a total reward or payoff function $p(\cdot)$. Formally, we consider the following optimization problem for the investor,

$$(1) \quad \sup_{\pi \in \mathcal{U}_{\mathbb{F}}} \mathbf{E}_{\pi}[p(\mathbf{x}, \mathbf{f}, \mathbf{r})],$$

where the real-valued reward function $p(\cdot)$ is a function of the entire sample path of portfolio positions \mathbf{x} , the factor realization \mathbf{f} , and security returns \mathbf{r} . For example, $p(\cdot)$ may have the form

$$(2) \quad p(\mathbf{x}, \mathbf{f}, \mathbf{r}) \triangleq W(\mathbf{x}, \mathbf{r}) - \text{TC}(\mathbf{u}) - \text{RA}(\mathbf{x}, \mathbf{f}, \mathbf{r}).$$

Here, W denotes the terminal wealth (total trading gains ignoring of transaction costs), i.e.,

$$(3) \quad W(\mathbf{x}, \mathbf{r}) \triangleq W_0 + \sum_{t=1}^T x_t^{\top} r_{t+1},$$

where W_0 is the initial wealth. $\text{TC}(\cdot)$ captures the transaction costs associated with a set of trading decisions, and $\text{RA}(\cdot)$ is the penalty term that incorporates risk aversion.

We make the following assumption about our objective function:

Assumption 3 (Concave objective function). *Given arbitrary, fixed sample paths of factor realizations \mathbf{f} and security returns \mathbf{r} , assume that the reward function $p(\mathbf{x}, \mathbf{f}, \mathbf{r})$ is a concave function of the sequence of positions \mathbf{x} .*

If $p(\cdot)$ has the specified form in (2), then Assumption 3 will be satisfied when the transaction cost term $\text{TC}(\cdot)$ is a convex function of trades and the risk aversion term $\text{RA}(\cdot)$ is a convex function of portfolio positions.

2.1. Examples

In this paper, we consider dynamic portfolio choice models that satisfy Assumptions 1–3. In order to illustrate the generality of this setting, we will now provide a number of specific examples that satisfy these assumptions.

In many cases, it may be more natural to model the percentage returns associated with an asset, rather than nominal price changes. Our framework accommodates such models, as we see in the following example:

Example 1 (Models of asset returns). *Consider an asset with price P_t , and with log-returns evolving according to*

$$\log\left(\frac{P_{t+1}}{P_t}\right) = g(F_t, \epsilon_{t+1}^{(1)}).$$

Here, F_t is a vector of predictive variables and $\epsilon_{t+1}^{(1)}$ is an i.i.d. disturbance term. We will assume that F_t is a Markov process, i.e.,

$$F_{t+1} = h(F_t, \epsilon_{t+1}^{(2)}),$$

where $\epsilon_{t+1}^{(2)}$ is another i.i.d. disturbance term.

In this setting, we can define the “factor” process $f_t \triangleq (P_t, P_{t-1}, F_t)$. This process evolves according to

$$f_{t+1} = G_{t+1}(f_t, \epsilon_{t+1}) \triangleq \left(P_t e^{g(F_t, \epsilon_{t+1}^{(1)})}, P_t, h(F_t, \epsilon_{t+1}^{(2)}) \right),$$

where $\epsilon_t \triangleq (\epsilon_t^{(1)}, \epsilon_t^{(2)})$. Similarly, define the price change process to be $r_t \triangleq P_t - P_{t-1}$. We have that

$$r_{t+1} = H_{t+1}(f_t, \epsilon_{t+1}) \triangleq P_t e^{g(F_t, \epsilon_{t+1}^{(1)})} - P_t,$$

Then, the joint dynamics of (f_t, r_t) satisfy Assumption 1.

Note that the Markovian assumption on the predictive variables in Example 1 is just for notational convenience and is not strictly necessary — we can always augment the vector with sufficient history so that the process becomes Markov. What is necessary is only that F_t be measurable with respect to the filtration generated by the disturbance processes. Indeed, the only real restriction that Assumption 1 imposes is that asset prices are exogenous and are not influenced by trades.

Example 2 (Gârleanu and Pedersen 2013). *This model has the following dynamics, where returns are driven by mean-reverting factors, that fit into our general framework:*

$$f_{t+1} = (I - \Phi) f_t + \epsilon_{t+1}^{(1)}, \quad r_{t+1} = \mu_t + B f_t + \epsilon_{t+1}^{(2)},$$

for each time $t \geq 0$. Here, μ_t is the deterministic ‘fair return’, e.g., derived from the CAPM, while $B \in \mathbb{R}^{N \times K}$ is a matrix of constant factor loadings. The factor process f_t is a vector mean-reverting process, with $\Phi \in \mathbb{R}^{K \times K}$ a matrix of mean reversion coefficients for the factors. It is assumed that the i.i.d. disturbances $\epsilon_{t+1} \triangleq (\epsilon_{t+1}^{(1)}, \epsilon_{t+1}^{(2)})$ are zero-mean with covariance given by $\text{Var}(\epsilon_{t+1}^{(1)}) = \Psi$ and $\text{Var}(\epsilon_{t+1}^{(2)}) = \Sigma$.

Trading is costly, and the transaction cost to execute $u_t = x_t - x_{t-1}$ shares is given by $\text{TC}_t(u_t) \triangleq \frac{1}{2} u_t^\top \Lambda u_t$, where $\Lambda \in \mathbb{R}^{N \times N}$ is a positive semi-definite matrix that measures the level of trading costs. There are no trading constraints (i.e., $U \triangleq \mathbb{R}^{N \times T}$). The investor’s objective function is to choose a trading strategy to maximize discounted future expected excess return, while accounting for transaction costs and adding a per-period penalty for risk, i.e.,

$$(4) \quad \underset{\pi \in \mathcal{U}_\pi}{\text{maximize}} \quad \mathbb{E}_\pi \left[\sum_{t=1}^T \left(x_t^\top B f_t - \text{TC}_t(u_t) - \text{RA}_t(x_t) \right) \right],$$

where $\text{RA}_t(x_t) \triangleq \frac{\gamma}{2} x_t^\top \Sigma x_t$ is a per-period risk aversion penalty, with γ being a coefficient of risk aversion. Gârleanu and Pedersen (2013) suggest this objective function for an investor who is compensated based on his performance relative to a benchmark. Each $x_t^\top B f_t$ term measures the excess return over the benchmark, while each $\text{RA}_t(x_t)$ term measures the variance of the tracking error relative to the benchmark.⁴

The problem (4) clearly falls into our framework. The objective function is similar to that of (2) with the minor variation that expected excess return rather than expected wealth is considered. Further, (4) has the further special property that total transaction costs and penalty for risk aversion decompose over time:

$$\text{RA}(\mathbf{x}, \mathbf{f}, \mathbf{r}) \triangleq \sum_{t=1}^N \text{RA}_t(x_t), \quad \text{TC}(\mathbf{u}) \triangleq \sum_{t=1}^N \text{TC}_t(u_t).$$

Note that this problem can be handled easily using the classical theory from the linear-quadratic control (LQC) literature (see, e.g., Bertsekas, 2000). This theory provides analytical characterization of optimal solution, for example, that the value function at any time t is quadratic function the state (x_t, f_t) , and that the optimal trade at each time is an affine function of the state. Moreover, efficient computational procedures are available to solve for the optimal policy.

On the other hand, the tractability of this model rests critically on three key requirements:

- The state variables (x_t, f_t) at each time t must evolve as linear functions of the control u_t and the i.i.d. disturbances ϵ_t (i.e., linear dynamics).

⁴See Gârleanu and Pedersen (2013) for other interpretations.

- Each control decision u_t is unconstrained.
- The objective function must decompose across time into a positive definite quadratic function of (x_t, u_t) at each time t .

These requirements are *not* satisfied by many real world examples, which may involve portfolio position or trade constraints, different forms of transaction costs and risk measures, and more complicated return dynamics. In the following examples, we will provide concrete examples of many such cases that do not admit optimal solutions via the LQC methodology, but remain within our framework.

Example 3 (Portfolio or trade constraints). *In practice, a common constraint in constructing equity portfolios is the short-sale restriction. Most of the mutual funds are enforced not to have any short positions by law. This requires the portfolio optimization problem to include the linear constraint*

$$x_t = x_0 + \sum_{s=1}^t u_s \geq \mathbf{0},$$

for each t . This is clearly a convex constraint on the set of feasible trade sequence \mathbf{u} .

We observe a similar restriction when an execution desk needs to sell or buy a large portfolio on behalf of an investor. Due to the regulatory rules in agency trading, the execution desk is only allowed to sell or buy during the trading horizon. In the ‘pure-sell’ scenario, the execution desk needs to impose the negativity constraint

$$u_t \leq \mathbf{0},$$

for each time t .

A third case arises in the context of insurance companies and banks that often need to satisfy certain minimum capital requirements in order to reduce the risk of insolvency. Therefore, they need to choose a dynamic investment portfolio so that their total wealth net of transaction costs exceeds a certain threshold C at all times. In our framework, this translates into a constraint

$$W_0 + \sum_{s=1}^t \left(x_s^\top r_{s+1} - \text{TC}_s(u_s) \right) \geq C,$$

for each time t and for each possible realization of returns \mathbf{r} . If each transaction cost function $\text{TC}_s(\cdot)$ is a convex function, then this constraint is also convex.

Each of the above well-known constraints in portfolio construction fit easily in our framework, but cannot be addressed via traditional LQC methods.

Example 4 (Non-quadratic transaction costs). *In practice, many trading costs such as the bid-ask spread, broker commissions, and exchange fees are intrinsically proportional to the trade size. Letting χ_i be the proportional transaction cost rate (an aggregate sum of bid-ask cost and commission fees, for example) for trading security i , the investor will incur a total cost of*

$$\text{TC}(\mathbf{u}) \triangleq \sum_{t=1}^T \sum_{i=1}^N \chi_i |u_{i,t}|.$$

The proportional transaction costs are a classical cost structure that is well studied in the literature (see, e.g., Constantinides, 1986).

Furthermore, other trading costs occur due to disadvantageous transaction price caused by the price impact of the trade. The management of the trading costs due to price impact has recently attracted considerable interest (see, e.g., Obizhaeva and Wang, 2005; Almgren and Chriss, 2000). Many models of transaction costs due to price impact imply a nonlinear relationship between trade size and the resulting transaction cost, for example

$$\text{TC}(\mathbf{u}) \triangleq \sum_{t=1}^T \sum_{i=1}^N \chi_i |u_{i,t}|^\beta.$$

Here, $\beta \geq 1$ and χ_i is a security-specific proportionality constant⁵.

In general, when the trade size is small relative to the total traded volume, proportional costs will dominate. On the other hand, when the trade size is large, costs due to price impact will dominate. Hence, both of these types of trading are important. However, the LQC framework of Example 2 only allows quadratic transaction costs (i.e., $\beta = 2$).

Example 5 (Terminal wealth risk). *The objective function of Example 2 includes a term to penalize excessive risk. In particular, the per-period quadratic penalty, $x_t^\top \Sigma x_t$, is used, in order to satisfy the requirements of the LQC model. However, penalizing risk additively in a per-period fashion is nonstandard. Such a risk penalty does not correspond to traditional forms of investor risk preferences, e.g., maximizing the expected utility of terminal wealth, and the economic meaning of such a penalty is not clear. An investor is typically more interested in the risk associated with the terminal wealth, rather than a sum of per-period penalties.*

In order to account for terminal wealth risk, let $\rho: \mathbb{R} \rightarrow \mathbb{R}$ be a real-valued convex function meant to penalize for excessive risk of terminal wealth (e.g., $\rho(w) = \frac{1}{2}w^2$ for a quadratic

⁵Gatheral (2010) notes that $\beta = \frac{3}{2}$ is a typical assumption in practice.

penalty) and consider the optimization problem

$$(5) \quad \underset{\pi \in \mathcal{U}_{\mathbb{F}}}{\text{maximize}} \quad \mathbb{E}_{\pi} \left[W(\mathbf{x}, \mathbf{r}) - \text{TC}(\mathbf{u}) - \gamma \rho(W(\mathbf{x}, \mathbf{r})) \right],$$

where $\gamma > 0$ is a risk-proportionality constant.

It is not difficult to see that the objective in (5) satisfies Assumption 3 and hence fits into our model. However, even when the risk penalty function $\rho(\cdot)$ is quadratic, (5) does *not* admit a tractable LQC solution, since the quadratic objective does not decompose across time.

Example 6 (Expected utility of terminal wealth). *Suppose that $U: \mathbb{R} \rightarrow \mathbb{R}$ is an increasing and concave utility function, and consider the optimization problem*

$$(6) \quad \underset{\pi \in \mathcal{U}_{\mathbb{F}}}{\text{maximize}} \quad \mathbb{E}_{\pi} \left[U(W(\mathbf{x}, \mathbf{r}) - \text{TC}(\mathbf{u})) \right].$$

Here, the objective is to maximize the expected utility of terminal wealth net of transaction costs. If the transaction cost function $\text{TC}(\cdot)$ is convex, the objective in (6) is the composition of a concave and increasing function and a concave function of \mathbf{x} ; this will be concave and satisfy Assumption 3.

Note that other mechanisms for risk aversion, such as penalties based on convex or coherent risk measures, can easily be incorporated in our framework in a manner analogous to Examples 5 and 6.

Example 7 (Maximum drawdown risk). *In addition to the terminal measures of risk described in Example 5, an investor might also be interested in controlling intertemporal measures of risk defined over the entire time trajectory. For example, a fund manager might be sensitive to a string of successive losses that may lead to the withdrawal of assets under management. One way to limit such losses is to control the maximum drawdown, defined as the worst loss of the portfolio between any two points of time during the investment horizon⁶. Formally,*

$$\text{MD}(\mathbf{x}, \mathbf{r}) \triangleq \max_{1 \leq t_1 \leq t_2 \leq T} \left(- \sum_{t=t_1}^{t_2} x_t^{\top} r_{t+1}, 0 \right).$$

It is easy to see that the maximum drawdown is a convex function of \mathbf{x} . Hence, the portfolio optimization problem

$$(7) \quad \underset{\pi \in \mathcal{U}_{\mathbb{F}}}{\text{maximize}} \quad \mathbb{E}_{\pi} \left[W(\mathbf{x}, \mathbf{r}) - \text{TC}(\mathbf{u}) - \gamma \text{MD}(\mathbf{x}, \mathbf{r}) \right],$$

⁶For example, see Grossman and Zhou (1993) for an earlier example.

where $\gamma \geq 0$ is a constant controlling trade-off between wealth and the maximum drawdown penalty, satisfies Assumption 3. Moreover, standard convex optimization theory yields that the problem (7) is equivalent to solving the constrained problem

$$(8) \quad \begin{aligned} & \underset{\pi \in \mathcal{U}_\pi}{\text{maximize}} && \mathbb{E}_\pi [W(\mathbf{x}, \mathbf{r}) - \text{TC}(\mathbf{u})] \\ & \text{subject to} && \mathbb{E}_\pi [\text{MD}(\mathbf{x}, \mathbf{r})] \leq C, \end{aligned}$$

where C (which depends on the choice of γ) is a limit on the allowed expected maximum drawdown.

Example 8 (Complex dynamics). We can also generalize the dynamics of Example 2. Consider factor and return dynamics given by

$$f_{t+1} = (I - \Phi) f_t + \epsilon_{t+1}^{(1)}, \quad r_{t+1} = \mu_t + (B + \xi_{t+1}) f_t + \epsilon_{t+1}^{(2)},$$

for each time $t \geq 0$. Here, each $\xi_{t+1} \in \mathbb{R}^{N \times K}$ is an extra noise term which captures model uncertainty regarding the factor loadings. We assume that

$$\mathbb{E}[(B + \xi_{t+1}) f_t \mid \mathcal{F}_t] = B f_t, \quad \text{Var}[(B + \xi_{t+1}) f_t \mid \mathcal{F}_t] = \bar{f}_t^\top \Upsilon \bar{f}_t,$$

where \mathcal{F}_t is the sigma-algebra incorporating all random variables realized by time t , and $\bar{f}_t \in \mathbb{R}^{K \times N}$ is a matrix given by $\bar{f}_t \triangleq [f_t \ f_t \ \dots \ f_t]$.

With this model, the conditional variance of returns becomes dependent on the factor structure and is time-varying, i.e., $\text{Var}[r_{t+1} \mid \mathcal{F}_t] = \bar{f}_t^\top \Upsilon \bar{f}_t + \Sigma$. This is consistent with the empirical work of Fama and French (1996), for example. In this setting, a per-period conditional variance risk penalty, analogous to that in (4) becomes $\text{RA}_t(\mathbf{x}, \mathbf{f}) = x_t^\top (\bar{f}_t^\top \Upsilon \bar{f}_t + \Sigma) x_t$. The resulting optimal control problem no longer falls into the LQC framework.

The dynamics and the reward functions considered in these examples satisfy our basic requirements of Assumptions 1–3. These examples illustrate that in many real-world problems with complex primitives for return predictability, transaction costs, risk measures and constraints, the dynamic portfolio choice becomes difficult to solve analytically or even using numerical methods when the number of assets is large.

3. Optimal Linear Model

The examples of Section 2.1 illustrated a broad range of important portfolio optimization problems. Without special restrictions, such as those imposed in the LQC framework, the

optimal dynamic policy for such a broad set of problems cannot be computed either analytically or computationally. In this section, in order to obtain policies in a computationally tractable way, we will consider a more modest goal. Instead of finding the optimal policy among *all* admissible dynamic policies, we will restrict our search to a subset of policies that are parsimoniously parameterized. That is, instead of solving for a globally optimal policy, we will instead find an approximately optimal policy by finding the best policy over the restricted subset of policies.

In order to simplify, we will assume that investor's reward function in (1) only depends on the sample path of portfolio positions \mathbf{x} and of factor realizations \mathbf{f} , and does not depend on the security returns \mathbf{r} explicitly. In other words, we assume that the reward function takes the form $p(\mathbf{x}, \mathbf{f})$. This is without loss of generality — given our general specification for factors under Assumption 1, we can simply include each security return as a factor. With this assumption, investor's trading decisions will, in general, be a non-anticipating function of the sample path of factor realizations \mathbf{f} . However, consider the following restricted set of policies, *linear rebalancing policies*, which are obtained by taking the affine combinations of the factors:

Definition 1 (Linear rebalancing policy). *A linear rebalancing policy π is a non-anticipating policy parameterized by a collection of vectors $\mathbf{c} \triangleq \{c_t \in \mathbb{R}^N, 1 \leq t \leq T\}$ and a collection of matrices $\mathbf{E} \triangleq \{E_{s,t} \in \mathbb{R}^{N \times K}, 1 \leq s \leq t \leq T\}$, that generates a sample path of trades $\mathbf{u} \triangleq (u_1, \dots, u_T)$ according to*

$$(9) \quad u_t \triangleq c_t + \sum_{s=1}^t E_{s,t} f_s,$$

for each time $t = 1, 2, \dots, T$.

Define \mathcal{C} to be the set of parameters (\mathbf{E}, \mathbf{c}) such that the resulting sequence of trades \mathbf{u} is contained in the constraint set U , with probability 1, i.e., \mathbf{u} is feasible. Denote by $\mathcal{L} \subset \mathcal{U}_{\mathbb{F}}$ the corresponding set of feasible linear policies.

Observe that linear rebalancing rules allow recourse, albeit in a restricted functional form. The affine specification (9) includes several classes of policies of particular interest as special cases:

- **Deterministic policies.** By taking $E_{s,t} \triangleq \mathbf{0}$, for all $1 \leq s \leq t \leq T$, it is easy to see that any deterministic policy is a linear rebalancing policy.
- **LQC optimal policies.** Optimal portfolios for the LQC framework of Example 2 take the form $x_t = \Gamma_{x,t} x_{t-1} + \Gamma_{f,t} f_t$, given matrices $\Gamma_{x,t} \in \mathbb{R}^{N \times N}$, $\Gamma_{f,t} \in \mathbb{R}^{N \times K}$, for all

$1 \leq t \leq T$, i.e., the optimal portfolio are linear in the previous position and the current factor values. Equivalently, by induction on t ,

$$x_t = \left(\prod_{s=1}^t \Gamma_{x,s} \right) x_0 + \sum_{s=1}^t \left(\prod_{\ell=1}^{s-1} \Gamma_{x,\ell} \right) \Gamma_{f,s} f_s.$$

Since $u_t = x_t - x_{t-1}$, it is clear that the optimal trade u_t is a linear function of the fixed initial position x_0 , and the factor realizations $\{f_1, \dots, f_t\}$, and is therefore of the form (9).

- **Linear portfolio policies.** Brandt et al. (2009) suggest a class of policies where portfolios are determined by adjusting a deterministic benchmark portfolio according to a linear function of a vector of stochastic, time-varying firm characteristics. In our setting, the firm characteristics would be interpreted as stochastic return predicting factors. An analogous rule would determine the positions at each time t via $x_t = \bar{x}_t + \Theta_t^\top (f_t - \bar{f}_t)$. Here, \bar{f}_t is the expected factor realization at time t . The policy is parameterized by \bar{x}_t , the deterministic benchmark portfolio at time t , and the matrix $\Theta_t \in \mathbb{R}^{N \times K}$, which maps firm characteristics (standardized to be mean zero) to adjustments to the benchmark portfolio. Such a portfolio rule is clearly of the form (9).
- **Policies based on basis functions.** Instead of having policies that are directly affine function of factor realizations, it is also possible to introduce basis functions (Skaf and Boyd, 2009). One might consider, for example, $\varphi: \mathbb{R}^K \rightarrow \mathbb{R}^D$, a collection of D (nonlinear) functions that capture particular features of the factor space that are important for good decision making. Consider a class of policies of the form

$$u_t \triangleq c_t + \sum_{s=1}^t E_{s,t} \varphi(f_s).$$

Such policies belong to the linear rebalancing class, if the factors are augmented also to include the value of the basis functions. This is easily accommodated in our framework, given the flexibility of Assumption 1. Similarly, policies which depend on the past security returns (in addition to factor realizations) can be accommodated by augmenting the factors with past returns.

- **Policies based on other policies.** One source of basis functions might be existing heuristic portfolio policies. For example, assume a collection of heuristic policies is available, each of which maps the history of factor realizations into a trading decision at each time. Each such map can be used to define a set of basis functions, as above.

The corresponding set of linear rebalancing policies would consist of all policies that are linear combinations of the heuristic policies.

An alternative to solving the original optimal control problem (1) is to consider the problem

$$(10) \quad \sup_{\pi \in \mathcal{L}} \mathbf{E}_{\pi}[p(\mathbf{x}, \mathbf{f})],$$

restricted to linear rebalancing rules. In general, (10) will not yield an optimal control for (1). The exception is if the optimal control for the problem is indeed a linear rebalancing rule (e.g., in a LQC problem). However, (10) will yield the optimal *linear* rebalancing rule. Further, in contrast to the original optimal control problem, (10) has the great advantage of being *tractable*, as suggested by the following result:

Proposition 1. *The optimization problem given by*

$$(11) \quad \begin{aligned} & \underset{\mathbf{E}, \mathbf{c}}{\text{maximize}} && \mathbf{E}[p(\mathbf{x}, \mathbf{f})] \\ & \text{subject to} && x_t = x_{t-1} + u_t, \quad \forall 1 \leq t \leq T, \\ & && u_t = c_t + \sum_{s=1}^t E_{s,t} f_s, \quad \forall 1 \leq t \leq T, \\ & && (\mathbf{E}, \mathbf{c}) \in \mathcal{C}. \end{aligned}$$

is a convex optimization problem, i.e., it involves the maximization of a concave function subject to convex constraints.

Proof. Note that $p(\cdot, \mathbf{f})$ is concave for a constant \mathbf{f} by Assumption 3. Since \mathbf{x} can be written as an affine transformation of (\mathbf{E}, \mathbf{c}) , then, for each fixed \mathbf{f} , the objective function is concave in (\mathbf{E}, \mathbf{c}) . Taking an expectation over realizations of \mathbf{f} preserves this concavity. Finally, the convexity of the constraint set \mathcal{C} follows from the convexity of U , under Assumption 2. ■

The problem (11) is a finite-dimensional, convex optimization problem that will yield parameters for the optimal linear rebalancing policy. It is also a stochastic optimization problem, in the sense that the objective is the expectation of a random quantity. In general, there are a number of effective numerical methods that can be applied to solve such problems:

- **Efficient exact formulation.** In many cases, with further assumptions on the problem primitives (the reward function $p(\cdot)$, the dynamics of the factor realizations \mathbf{f} , and the trading constraint set U), the objective $\mathbf{E}[p(\mathbf{x}, \mathbf{f})]$ and the constraint set \mathcal{C} of

the program (11) can be analytically expressed explicitly in terms of the decision variables (\mathbf{E}, \mathbf{c}) . In some of these cases, the program (11) can be transformed into a standard form of convex optimization program such as a quadratic program or a second-order cone program. In such cases, off-the-shelf solvers specialized to these standard forms (e.g., Grant and Boyd, 2011) can be used. Alternatively, generic methods for constrained convex optimization such as interior point methods (see, e.g., Boyd and Vandenberghe, 2004) can be applied to efficiently solve large-scale instances of (11). We will explore this topic further, developing a number of efficient exact formulations in Appendix A, and providing numerical examples in Sections 4–5.

- **Sample average approximation (SAA).** In the absence of further structure on the problem primitives, the program (11) can also be solved via Monte Carlo sampling. Specifically, suppose that $\mathbf{f}^{(1)}, \dots, \mathbf{f}^{(S)}$ are S independent sample paths of factor realization. The objective and constraints of (11) can be replaced with sampled versions, to obtain

$$\begin{aligned}
 (12) \quad & \underset{\mathbf{E}, \mathbf{c}}{\text{maximize}} && \frac{1}{S} \sum_{\ell=1}^S p(\mathbf{x}^{(\ell)}, \mathbf{f}^{(\ell)}) \\
 & \text{subject to} && x_t^{(\ell)} = x_{t-1}^{(\ell)} + u_t^{(\ell)}, \quad \forall 1 \leq t \leq T, 1 \leq \ell \leq S, \\
 & && u_t^{(\ell)} = c_t + \sum_{s=1}^t E_{s,t} f_s^{(\ell)}, \quad \forall 1 \leq t \leq T, 1 \leq \ell \leq S, \\
 & && \mathbf{u}^{(\ell)} \in U, \quad \forall 1 \leq \ell \leq S.
 \end{aligned}$$

The sample average approximation (12) can be solved via standard convex optimization methods (e.g., interior point methods). Moreover, under appropriate regularity conditions, convergence of the SAA (12) to the original program (11) can be established as $S \rightarrow \infty$, along with guarantees on the rate of convergence (Shapiro, 2003).

- **Stochastic approximation.** Denote the collection of decision variables in (11) by $\mathbf{z} \triangleq (\mathbf{E}, \mathbf{c})$, and, allowing a minor abuse of notation, define $p(\mathbf{z}, \mathbf{f})$ to be the reward when the sample path of factor realizations is given by \mathbf{f} and the trading policy is determined by \mathbf{z} . Then, defining $h(\mathbf{z}) \triangleq p(\mathbf{z}, \mathbf{f})$, the problem in (11) is simply to maximize $\mathbb{E}[h(\mathbf{z})]$ subject to the constraint that $\mathbf{z} \in \mathcal{C}$. Under suitable technical conditions, super-differentials of h and p are related according to $\partial h(\mathbf{z}) = \mathbb{E}[\partial_{\mathbf{z}} p(\mathbf{z}, \mathbf{f})]$. Stochastic approximation methods are incremental methods that seek to estimate ascent directions for $h(\cdot)$ from sampled ascent directions for $p(\cdot, \mathbf{f})$. For example, given a sequence of i.i.d. sample paths of factor realizations $\mathbf{f}^{(1)}, \mathbf{f}^{(2)}, \dots$, a sequence of parameter esti-

mates $\mathbf{z}^{(1)}, \mathbf{z}^{(2)}, \dots$ can be constructed according to

$$\mathbf{z}^{(\ell+1)} = \Pi_{\mathcal{C}} \left(\mathbf{z}^{(\ell)} + \gamma_{\ell} \zeta_{\ell} \right),$$

where $\Pi_{\mathcal{C}}(\cdot)$ is the projection onto the feasible set \mathcal{C} , $\zeta_{\ell} \in \partial_{\mathbf{z}p}(\mathbf{z}^{(\ell)}, \mathbf{f}^{(\ell)})$ is a supergradient, and $\gamma_{\ell} > 0$ is a step-size. Stochastic approximation methods have the advantage of being incremental and thus requiring minimal memory relative to sample average approximation, and are routinely applied in large scale convex stochastic optimization (Nemirovski et al., 2009).

One attractive feature of our framework is that it often can be applied in a *data-driven* fashion, without separately specifying and estimating an explicit functional form for the factor and return dynamics. For example, the sample average approximation and stochastic approximation approaches only need access to simulated trajectories of factors and returns — they do not need explicit knowledge of the dynamics in Assumption 1 that drive these processes. It may be possible to use historical factor and return realizations (possibly in combination with non-parametric methods such as bootstrapping) to generate sample trajectories without an explicit model of the underlying dynamics. Similarly, in many of the exact formulations developed in Appendix A, including the numerical examples of Sections 4–5, only moments of the factor realizations are necessary in order to find the optimal linear rebalancing policy. These can be estimated from historical data without an explicit, calibrated model.

Finally, observe that optimal linear policies can also be applied in concert with *model predictive control* (MPC). Here, at each time step t , the program (11) is resolved beginning from time t . This determines the optimal linear rebalancing rule from time t forward, conditioned on the realized history up to time t . The resulting policy is only used to determine the current trading decision at time t , and (11) is subsequently resolved at each future time period. At the cost of an additional computational burden, the use of optimal linear policies with MPC subsumes standard MPC approaches, such as resolving a myopic variation of the portfolio optimization problem (and ignoring the true multi-period nature) or solving a deterministic variation of the portfolio optimization problem (and ignoring the possibility of future recourse).

4. Application: Equity Agency Trading

In this section, we provide an empirical application to illustrate the implementation and the benefits of the optimal linear policy. As our example, we consider an important problem

in equity agency trading. Equity agency trading seeks to address the problem faced by large investors such as pension funds, mutual funds, or hedge funds that need to update the holdings of large portfolios. Here, the investor seeks to minimize the trading costs associated with a large portfolio adjustment. These costs, often labeled ‘execution costs’, consist of commissions, bid-ask spreads, and, most importantly in the case of large trades, price impact from trading. Efficient execution of large trades is accomplished via ‘algorithmic trading’, and requires significant technical expertise and infrastructure. For this reason, large investors utilize algorithmic trading service providers, such as execution desks in investment banks. Such services are often provided on an agency basis, where the execution desk trades on behalf of the client, in exchange for a fee. The responsibility of the execution desk is to find a feasible execution schedule over the client-specified trading horizon while minimizing trading costs and aligning with the risk objectives of the client.

The problem of finding an optimal execution schedule has received a lot of attention in the literature since the initial paper of Bertsimas and Lo (1998). In their model, when price impact is proportional to the number of shares traded, the optimal execution schedule is to trade equal number of shares at each trading time. There are number of papers that extend this model to incorporate the risk of the execution strategy. For example, Almgren and Chriss (2000) derive that risk averse agents need to liquidate their portfolio faster in order to reduce the uncertainty of the execution cost.

The models described above seek mainly to minimize execution costs by accounting for the price impact and supply/demand imbalances caused by the investor’s trading. Complementary to this, an investor may also seek to exploit short-term predictability of stock returns to inform the design of a trade schedule. As such, there is a growing interest to model return predictability in intraday stock returns. Often called ‘short-term alpha models’, some of the predictive models are similar to well-known factor models for the study of long-term stock returns, e.g., the Capital Asset Pricing Model (CAPM), or the Fama-French Three Factor Model. Alternatively, short-term predictions can be developed from microstructure effects, for example the imbalance of orders in an electronic limit order book. Heston et al. (2010) document that systematic trading as described in the examples above and institutional fund flows lead to predictable patterns in intraday returns of common stocks.

We will consider an agency trading optimal execution problem in the presence of short-term predictability. One issue that arises here is that, due to the regulatory rules in agency trading, the execution desk is only allowed to either sell or buy a particular security over the course of the trading horizon, depending on whether the ultimate position adjustment desired for that security is negative or positive. However, given a model for short-term predictability, an optimal trading policy that minimizes execution cost may result in both

buy and sell trades for the same security as it seeks to exploit short-term signals. Hence, it is necessary to impose constraints on the sign of trades, as in Example 3.

If an agency trading execution problem has price and factor dynamics which satisfy Assumption 1 and an objective (including transaction costs, price impact, and risk aversion) that satisfies Assumption 3, then we can compute the best execution schedule in the space of linear execution schedules, i.e., the number of shares to trade at each time is a linear function of the previous return predicting factors. We will consider a particular formulation that involves linear price and factor dynamics and a quadratic objective function (as in Example 2). Note that this example does not highlight the full generality of our framework — more interesting cases would involve nonlinear factor dynamics (e.g., microstructure-based order imbalance signals) or a non-quadratic objective (e.g., transaction costs as in Example 4). However, this example is intentionally chosen since, in the absence of the trade sign constraint, the problem can be solved exactly with LQC methods. Hence, we are able to compare the optimal linear policy to policies derived from LQC methods applied to the unconstrained problem.

The rest of this section is organized as follows. We present our optimal execution problem formulation in Section 4.1. An exact, analytical solution is not available to this problem, hence, in Section 4.2, we describe several approximate solution techniques, including finding the best linear policy. In order to evaluate the quality of the approximate methods, in Section 4.3, we describe several techniques for computing upper bounds on the performance of any policy for our execution problem. In Section 4.4, we describe the empirical calibration of the parameters of our problem. Finally, in Section 4.5, we present and discuss the numerical results.

4.1. Formulation

We follow the general framework of Section 2. Suppose that $x_0 \in \mathbb{R}^N$ denotes the number of shares in each of N securities that we would like to sell before time T . We assume that trades can occur at discrete times, $t = 1, \dots, T$. We define an *execution schedule* to be the collection $\mathbf{u} \triangleq (u_1, \dots, u_T)$, where each $u_t \in \mathbb{R}^N$ denotes the number of shares traded at time t . Note that a negative (positive) value of $u_{i,t}$ denotes a sell (buy) trade of security i at time t . The total position at time t is given by $x_t = x_0 + \sum_{s=1}^t u_s$.

The formulation of the agency trading optimal execution problem is as follows:

- **Constraints.** Without loss of generality, we will assume that the initial position is positive, i.e., $x_0 > \mathbf{0}$. The execution schedule must liquidate the entire initial position

by the end of the time horizon, thus

$$(13) \quad x_T = x_0 + \sum_{t=1}^T u_t = \mathbf{0}.$$

Further, agency trading regulations allow only sell trades, thus

$$(14) \quad u_t \leq \mathbf{0}, \quad t = 1, \dots, T.$$

Note that any schedule satisfying (13)–(14) will also satisfy

$$(15) \quad x_t = x_0 + \sum_{s=1}^t u_s \geq \mathbf{0}, \quad t = 1, \dots, T.$$

We denote by $\mathcal{U}_{\mathbb{F}}^0$ the set of non-anticipating policies satisfying (13) almost surely, and by $\mathcal{U}_{\mathbb{F}}$ the set of non-anticipating policies satisfying (13)–(15) almost surely.

- **Return and factor dynamics.** We follow the discrete time linear dynamics of Gârleanu and Pedersen (2013),⁷ as described in Example 2. We assume that the price change of each security from t to $t + 1$ is given by the vector r_{t+1} , and is predicted by K factors collected in a vector f_t . Furthermore, the evolution of factor realizations follow a mean reverting process. Formally, we have the following dynamics for price changes and factor realizations:

$$f_{t+1} = (I - \Phi) f_t + \epsilon_{t+1}^{(1)}, \quad r_{t+1} = \mu + B f_t + \epsilon_{t+1}^{(2)},$$

where $B \in \mathbb{R}^{N \times K}$ is a constant matrix of factor loadings, $\Phi \in \mathbb{R}^{K \times K}$ is a diagonal matrix of mean reversion coefficients for the factors, and $\mu \in \mathbb{R}^N$ is the mean return. We assume that the noise terms are i.i.d., and normally distributed with zero-mean and with covariance matrices given by $\text{Var}(\epsilon_{t+1}^{(1)}) = \Psi \in \mathbb{R}^{N \times N}$ and $\text{Var}(\epsilon_{t+1}^{(2)}) = \Sigma \in \mathbb{R}^{K \times K}$. We discuss the precise choice of return predicting factors and the calibration of the dynamics shortly in Section 4.4.

- **Objective.** We assume that the investor is risk-neutral and seeks to maximize total

⁷Note that Gârleanu and Pedersen (2013) consider an infinite horizon setting, while our setting is finite horizon. Further, Gârleanu and Pedersen (2013) solve for dynamic policies in the absence of the constraints (13)–(15).

excess profits after quadratic transaction costs, i.e.,

$$(16) \quad \mathcal{V}^* \triangleq \underset{\pi \in \mathcal{U}_{\mathbb{F}}}{\text{maximize}} \quad \mathbb{E}_{\pi} \left[\sum_{t=1}^T \left(x_t^{\top} B f_t - \frac{1}{2} u_t^{\top} \Lambda u_t \right) \right],$$

where $\Lambda \in \mathbb{R}^{N \times N}$ is a matrix parameterizing the quadratic transaction costs.

Note that the problem (16) is a special case of the optimization program in Example 2, with the exception of the constraints (13)–(15).

4.2. Approximate Policies

Since a tractable analytical or computational solution to the optimal execution problem in (16) is not available, we compare four approximate solution techniques:

- **TWAP.** A time-weighted average price (TWAP) policy seeks to sell a fixed quantity $u_t = -x_0/T$ of shares in each of the T periods. This policy minimizes transaction costs, and would be optimal in the absence of a predictive model for returns.
- **Deterministic.** Instead of allowing for a non-anticipating dynamic policy, where the trade at each time t is allowed to depend on all events that have occurred before t , we can solve for an optimal static policy, i.e., a deterministic sequence of trades over the entire time horizon that is decided at the beginning of the time horizon. Here, observe that at the beginning of the time horizon, the expected future factor vector is given by $\mathbb{E}[f_t|f_0] = (I - \Phi)^t f_0$. Therefore, in order to find the optimal deterministic policy, given f_0 , we maximize the conditional expected value of the stochastic objective in (16) by solving the quadratic program

$$(17) \quad \begin{aligned} & \underset{\mathbf{u}}{\text{maximize}} && \sum_{t=1}^T \left(x_t^{\top} B (I - \Phi)^t f_0 - \frac{1}{2} u_t^{\top} \Lambda u_t \right) \\ & \text{subject to} && u_t = x_t - x_{t-1}, && t = 1, \dots, T, \\ & && u_t \leq \mathbf{0}, \quad x_t \geq \mathbf{0}, && t = 1, \dots, T, \\ & && x_T = \mathbf{0}, \end{aligned}$$

to yield a deterministic sequence of trades \mathbf{u} .

- **Model predictive control.** In this approximation, at each trading time, we solve for the deterministic sequence of trades conditional on the available information and implement only the first trade. Thus, this policy is an immediate extension of the

deterministic policy, with the addition of resolving at each trading time. Formally, at time t , we solve the quadratic program

$$\begin{aligned}
(18) \quad & \underset{u_t, \dots, u_T}{\text{maximize}} && \sum_{s=t}^T \left(x_s^\top B (I - \Phi)^{(s-t)} f_t - \frac{1}{2} u_s^\top \Lambda u_s \right) \\
& \text{subject to} && u_s = x_s - x_{s-1}, && s = t, \dots, T, \\
& && u_s \leq \mathbf{0}, \quad x_s \geq \mathbf{0}, && s = t, \dots, T, \\
& && x_T = \mathbf{0}.
\end{aligned}$$

If (u_t^*, \dots, u_T^*) is the optimal solution, then the investor trades u_t^* at time t .

- **Projected LQC.** If the inequality constraints (14)–(15) are eliminated, the program would reduce to the classical linear quadratic control problem

$$(19) \quad \underset{\pi \in \mathcal{U}_{\mathbb{F}}^0}{\text{maximize}} \quad \mathbf{E}_\pi \left[\sum_{t=1}^T \left(x_t^\top B f_t - \frac{1}{2} u_t^\top \Lambda u_t \right) \right].$$

The optimal dynamic policy for the program in (19) yields the trade

$$(20) \quad u_t = (\Lambda + A_{xx,t})^{-1} (\Lambda x_{t-1} + (B + A_{xf,t} (I - \Phi)) f_t) - x_{t-1}$$

at each time t as a function of the previous position x_{t-1} and the current factor values f_t . Here, the matrices $A_{xx,t}$ and $A_{xf,t}$ are derived in Appendix B. The dynamic rule for u_t in (20) of course will not be feasible for the constrained program in (16), in general. Thus, the projected LQC policy seeks a trade decision, \hat{u}_t , which is the projection of u_t onto the constraint set (14)–(15). In other words, given a trading decision $u_{i,t}$, we find the closest trade $\hat{u}_{i,t}$ among all trades satisfying (14)–(15), according to $\hat{u}_{i,t} = \max \{-x_{i,t-1}, \min \{0, u_{i,t}\}\}$, for each time $t < T$ and for each security i .

- **Optimal linear.** As formulated in Definition 1, a linear rebalancing policy specifies trades according to

$$u_t \triangleq c_t + \sum_{s=1}^t E_{s,t} f_s,$$

for each time $t = 1, 2, \dots, T$, given parameters (\mathbf{E}, \mathbf{c}) . Due to the linear relationship between position and trade vectors, we can represent the position vector in the similar form, i.e., $x_t = d_t + \sum_{s=1}^t J_{s,t} f_s$ where $d_t \triangleq x_0 + \sum_{i=1}^t c_i$ and $J_{s,t} \triangleq \sum_{i=s}^t E_{s,i}$. As shown in Appendix A.1.1, we implement the almost sure equality constraint (13) via equality constraints on the policy parameters by setting $d_T = \mathbf{0}$, and $J_{t,T} = \mathbf{0}$ for all t . We replace the almost sure inequality constraints (14)–(15) with probabilistic relaxations,

as discussed in Appendix A.1.2. With these assumptions, we compute the parameters of the optimal linear policy by solving the following stochastic program:

(21)

$$\begin{aligned}
& \underset{(\mathbf{E}, \mathbf{c})}{\text{maximize}} && \mathbb{E} \left[\sum_{t=1}^T \left(\left(d_t + \sum_{s=1}^t J_{s,t} f_s \right)^\top B f_t - \frac{1}{2} \left(c_t + \sum_{s=1}^t E_{s,t} f_s \right)^\top \Lambda \left(c_t + \sum_{s=1}^t E_{s,t} f_s \right) \right) \right] \\
& \text{subject to} && d_t = x_0 + \sum_{i=1}^t c_i, && 1 \leq t \leq T, \\
& && J_{s,t} = \sum_{i=s}^t E_{s,i}, && 1 \leq s \leq t \leq T, \\
& && \mathbb{P} \left(d_t + \sum_{s=1}^t J_{s,t} f_s < \mathbf{0} \right) \leq \eta, && 1 \leq t \leq T, \\
& && \mathbb{P} \left(c_t + \sum_{s=1}^t E_{s,t} f_s > \mathbf{0} \right) \leq \eta, && 1 \leq t \leq T, \\
& && d_T = \mathbf{0}, \\
& && J_{t,T} = \mathbf{0}, && 1 \leq t \leq T.
\end{aligned}$$

Here, the parameter $\eta \in (0, 1/2)$ controls the probability that the constraints (14)–(15) are violated.⁸ Using the fact that the objective is an expectation of a quadratic expression in Gaussian random variables and the fact that the chance constraints can be handled using Lemma 1 in Appendix A, (21) can be explicitly written as a second-order cone program. This calculation is detailed in Appendix C. Then, (21) can be solved using an off-the-shelf convex optimization solver.

The solution of (21) provides the desired linear policy, $u_t = c_t + \sum_{s=1}^t E_{s,t} f_s$, in the return predicting factors. However, due to the fact that some of the constraints of the original program in (16) are only probabilistically enforced, u_t may not be feasible for the original program. The projected optimal linear policy seeks a trade decision, \hat{u}_t , which is the projection of u_t onto the constraint set (14)–(15). In other words, given a trading decision $u_{i,t}$, we find the closest trade $\hat{u}_{i,t}$ among all trades satisfying (14)–(15), according to $\hat{u}_{i,t} = \max \{-x_{i,t-1}, \min \{0, u_{i,t}\}\}$, for each time $t < T$ and security i .

4.3. Upper Bounds

In order to evaluate the quality of the policies described in Section 4.2, we compute a number of upper bounds on the performance of *any* policy for the program (16), as follows:

- **Perfect foresight.** In this upper bound, we compute the value of an optimal policy with the perfect knowledge of future factor values. In particular, given a vector of

⁸We used the value $\eta = 0.2$ in our simulation results.

factor realizations \mathbf{f} , consider the optimization problem

$$\begin{aligned}
(22) \quad \mathcal{V}^{\text{PF}}(\mathbf{f}) \triangleq & \underset{\mathbf{u}}{\text{maximize}} \quad \sum_{t=1}^T \left(x_t^\top B f_t - \frac{1}{2} u_t^\top \Lambda u_t \right) \\
& \text{subject to} \quad u_t = x_t - x_{t-1}, & t = 1, \dots, T, \\
& u_t \leq \mathbf{0}, \quad x_t \geq \mathbf{0}, & t = 1, \dots, T, \\
& x_T = \mathbf{0}.
\end{aligned}$$

The value $\mathcal{V}^{\text{PF}}(\mathbf{f})$ is the best that can be achieved with perfect foresight of a particular sample path of factors \mathbf{f} . Note that this can be readily computed by solving the quadratic program (22). Since the non-anticipating policies of the original program (16) are not able to utilize future factor information in making trading decisions, we have the upper bound $\mathcal{V}^* \leq \mathbb{E}[\mathcal{V}^{\text{PF}}(\mathbf{f})]$. This upper bound can be computed via Monte Carlo simulation over sample paths of factor realizations.

- **Unconstrained LQC.** The value of the LQC problem (19), where the inequality constraints (14)–(15) are relaxed, also provides an upper bound to (16). The expected value of the relaxed program can be exactly computed and yields the upper bound

$$(23) \quad \mathcal{V}^* \leq -\frac{1}{2} x_0^\top A_{xx,0} x_0 + \frac{1}{2} \left(\text{tr} \left(\Omega_0 (I - \Phi)^\top A_{ff,0} (I - \Phi) \right) + \sum_{t=0}^{T-2} \text{tr}(\Psi A_{ff,t}) \right),$$

where the matrices $A_{xx,0}$ and $A_{ff,t}$ are derived in the Appendix B.

- **Pathwise optimization.** Given a sample path \mathbf{f} of factor realizations and a sequence $\boldsymbol{\zeta} \triangleq (\zeta_1, \dots, \zeta_T)$ of vectors $\zeta_t \in \mathbb{R}^K$ for each t , consider the quadratic optimization program

$$\begin{aligned}
(24) \quad \mathcal{V}^{\text{PO}}(\mathbf{f}, \boldsymbol{\zeta}) \triangleq & \underset{\mathbf{u}}{\text{maximize}} \quad \sum_{t=1}^T \left(x_t^\top B f_t - \zeta_t^\top \epsilon_t^{(1)} - \frac{1}{2} u_t^\top \Lambda u_t \right) \\
& \text{subject to} \quad \epsilon_t^{(1)} = f_t - (I - \Phi) f_{t-1}, & t = 1, \dots, T, \\
& u_t = x_t - x_{t-1}, & t = 1, \dots, T, \\
& u_t \leq \mathbf{0}, \quad x_t \geq \mathbf{0}, & t = 1, \dots, T, \\
& x_T = \mathbf{0}.
\end{aligned}$$

It can be established (Desai et al., 2011; Brown and Smith, 2010) that for any $\boldsymbol{\zeta}$, the upper bound $\mathcal{V}^* \leq \mathbb{E}[\mathcal{V}^{\text{PO}}(\mathbf{f}, \boldsymbol{\zeta})]$ holds — observe that the perfect foresight upper bound is a special case of this when $\boldsymbol{\zeta}$ is zero. Roughly speaking, this upper bound corresponds to a relaxation of the non-anticipating policy requirement, and $\boldsymbol{\zeta}$ correspond to a choice of Lagrange multipliers for this relaxation. The pathwise optimization upper bound

corresponds to making a choice for ζ that results in an *optimal* upper bound, i.e., $\mathcal{V}^* \leq \min_{\zeta} \mathbb{E}[\mathcal{V}^{\text{PO}}(\mathbf{f}, \zeta)]$. This minimization involves a convex objective function and can be computed via stochastic gradient descent; we refer the reader to Desai et al. (2011) for details.

4.4. Model Calibration

In this section, we describe calibration of the parameters of the optimal execution problem formulated in Section 4.1. We chose one of the most liquid stocks, Apple, Inc. (NASDAQ: AAPL), for our empirical study. We set the execution horizon to be 1 hour and trade intervals to be 5 minutes. Thus, setting a trade interval to be a one unit of time, we have a time horizon of $T = 12$. We assume that the initial position to be liquidated is $x_0 = 100,000$ shares.

In trade execution problems, the time horizon is typically a day, thus we will construct a factor model in the same time-frequency. We will use the intraday transaction prices of AAPL from the NYSE TAQ database on the trading days of January 4, 2010 (day 0) and January 5, 2010 (day 1) to construct $K = 2$ return predicting factors, each with a different mean reversion speed. We first divide each trading day into 78 time intervals, each 5 minutes in length. For each 5 minute interval, we calculate the average transaction price from all transactions in that interval. Let $p_t^{(d)}$ be the average price for interval $t = 1, \dots, 78$ on day $d = 0, 1$. Let $f_{k,t}$ be the value of factor $k = 1, 2$ for interval $t = 2, \dots, 78$, defined as follows

$$f_{1,t} \triangleq p_t^{(1)} - p_{t-1}^{(1)}, \quad f_{2,t} \triangleq p_t^{(1)} - p_t^{(0)}.$$

In other words, $f_{1,t}$ is the average price change over the previous 5 minute interval, while $f_{2,t}$ is the average price change relative to the previous day. Here, we can interpret the factors as the representations of value and momentum signals. Intuitively, the first factor can be considered as a ‘momentum’-type signal with fast mean reversion and the second factor as a ‘value’-type signal with slow mean reversion.

Given the price change of the security $r_{t+1} \triangleq p_{t+1}^{(1)} - p_t^{(1)}$, we can compute the estimate of the factor loading matrix, B , using the following linear regression:

$$r_{t+1} = 0.0726 + 0.3375 f_{1,t} - 0.0720 f_{2,t} + \epsilon_{t+1}^{(2)},$$

(1.96) (3.11) (-2.2)

where the OLS t-statistics are reported in brackets. Thus,

$$B = \begin{bmatrix} 0.3375 & -0.072 \end{bmatrix}.$$

Similarly, we obtain the mean reversion rates for the factors,

$$\begin{aligned} \Delta f_{1,t+1} &= -0.7146 f_{1,t} + \epsilon_{1,t+1}^{(1)}, & \Delta f_{2,t+1} &= -0.0353 f_{2,t} + \epsilon_{2,t+1}^{(1)}. \\ &(-6.62) & &(-1.16) \end{aligned}$$

Thus,

$$\Phi = \begin{bmatrix} 0.7146 & 0 \\ 0 & 0.0353 \end{bmatrix}.$$

The variance of the error terms is estimated to be

$$\Sigma \triangleq \text{Var}(\epsilon_t^{(2)}) = 0.0428, \quad \Psi \triangleq \text{Var}(\epsilon_t^{(1)}) = \begin{bmatrix} 0.0378 & 0 \\ 0 & 0.0947 \end{bmatrix}.$$

The distribution of the initial factor realization, f_0 , is set to the stationary distribution under the given factor dynamics, i.e., f_0 is normally distributed with zero mean and covariance

$$\Omega_0 \triangleq \sum_{t=1}^{\infty} (I - \Phi)^t \Psi (I - \Phi)^t = \begin{bmatrix} 0.0412 & 0 \\ 0 & 1.3655 \end{bmatrix}.$$

A rough estimate of the transaction cost coefficient $\Lambda = 2.14 \times 10^{-5}$ is used — this implies a transaction cost of \$10 or 0.5 basis points on a typical trade of 1,000 shares.

4.5. Numerical Results

Using the calibrated parameters from Section 4.4, we run a simulation with 50,000 trials to estimate the performance of each of the approximate policies of Section 4.2. In each trial, we sample the initial factor f_0 , solve for the resulting policy of each approximate method, and compute its corresponding payoff. In order to evaluate the performance of each policy effectively, we use the same set of simulation paths in each policy’s computation of average payoff. We used CVX (Grant and Boyd, 2011), a package for solving convex optimization problems in Matlab, to solve the optimization problems that occur in the computation of the deterministic, model predictive control, and optimal linear policies.

The upper half of Table 1 summarizes the performance of each of the policies described in Section 4.2. For each policy, we divide the total payoff into two components, the alpha gains (i.e., $\sum_{t=1}^T x_t^\top B f_t$) and the transaction costs (i.e., $\sum_{t=1}^T -u_t^\top \Lambda u_t$). For each of these components as well as the total, we report the mean value over all simulation trials and the associated standard error.⁹ In the lower half of Table 1, we report upper bounds on the

⁹Note that values for the TWAP policy and the unconstrained LQC upper bound are computed exactly without Monte Carlo simulation.

			Alpha (\$k)	T.C. (\$k)	Total (\$k)	Optimality Gap	CPU time (sec)
Policies	TWAP	Mean	0.03	-8.91	-8.88	238%	< 0.01
		S.E.	0.210	0.000	0.210		
	Deterministic	Mean	19.34	-15.81	3.53	45.4%	0.82
		S.E.	0.229	0.025	0.224		
	Model predictive control	Mean	21.25	-16.54	4.71	27.1%	5.79
		S.E.	0.233	0.023	0.225		
	Projected LQC	Mean	25.13	-19.40	5.73	11.3%	0.02
		S.E.	0.227	0.039	0.229		
	Optimal linear	Mean	23.24	-17.11	6.13	5.11%	4.23
		S.E.	0.233	0.025	0.224		
Upper Bounds	Pathwise optimization	Mean			6.46		
		S.E.			0.04		
	Perfect foresight	Mean				8.57	
		S.E.				0.223	
	Unconstrained LQC	Mean				12.58	
		S.E.					

Table 1: Summary of the performance statistics of each policy in the optimal execution example, along with upper bounds. In the upper half of the table we consider the approximate policies. For each approximate policy, we divide the total payoff into two components, the alpha gains and the transaction costs. For each performance statistic, we report the mean value and the associated standard error. Finally, we report the average computation time (in seconds) for each policy per simulation trial. In the bottom half of the table, we report the computed upper bounds on the total payoff. For those methods which involve Monte Carlo simulation, standard errors are also reported.

total payoff of any policy, as computed using the methods described in Section 4.3. The pathwise optimization method achieves the tightest upper bound. For each policy, we report an optimality gap relative to this tightest upper bound.

Comparing the performance of the various policies in Table 1, we see accounting for predictable price movements can make a significant difference. Indeed, the TWAP policy, which minimizes transaction costs but ignores predictable price movements, performs the worst. Other policies incur higher transaction costs than TWAP but more than make up for this by opportunistically timing the liquidation relative to predictable price movements. Of the remaining policies, the projected LQC and optimal linear policies achieve the highest performance. These are the only policies that are constructed in a manner that explicitly account for the dynamic multi-period nature of the problem and allow for recourse.

The overall best policy is the optimal linear policy, which achieves a value that is within 5% of the value that can be achieved by any policy. This optimality gap is a factor of

	(Optimal Linear) – (Projected LQC)		
	Alpha	T.C.	Total
	(\$k)	(\$k)	(\$k)
Mean	-1.89	2.29	0.40
S.E.	0.0137	0.0196	0.0095

Table 2: A detailed comparison of the difference in alpha gains, transaction costs, and total performance between the optimal linear policy and projected dynamic policy in the optimal execution example. We observe that the standard error for the difference in total payoff is very small, thus, the performance gain by employing the optimal linear policy is statistically significant.

two improvement over the optimality gap of the next best policy, projected LQC, and is significantly better than any other policy.

Note that, despite the higher total payoff for the optimal linear policy as compared to the projected LQC policy in Table 1, the relatively high standard errors preclude the immediate conclusion that the optimal linear policy achieves a statistically significant higher total payoff. Thus, in order to provide a more careful comparison, for each simulation trial, we consider the difference in alpha gains, transaction costs, and total payoff between these two policies. Table 2 show the statistics of these differences, and establishes that the performance benefit of the optimal linear policy is statistically significant. Moreover, Table 2 reveals that the optimal linear policy achieves a better result by more carefully managing transaction costs, at the expense of not achieving the alpha gains of the projected LQC policy.

In Table 1, we also report the average computation time (in seconds) required to evaluate each policy for over a single simulated sample path. This gives a sense of the relative computational complexity of the various policies. The TWAP and projected LQC policies are the fastest to evaluate — the former is essentially trivial, while the latter has a closed-form expression (via a solution of recursive equations). The remaining policies involve solving at least one optimization problem per sample path. These policies have roughly the same order of magnitude in computation time, with model predictive control (which solves a different optimization problem at every time step) having the longest running time.

4.6. Sensitivity Results

In Table 3, we report the sensitivity of our simulation results with respect to main parameters of the optimal execution problem, e.g., length of time horizon, level of transaction costs, level of factor persistence and relaxation probability. We only vary the parameter at hand while keeping the other parameters fixed. We report the average objective value and its standard

error for the optimal linear and projected LQC policies, the top performing policies in our baseline simulation.

We observe that optimal linear policy outperforms the projected LQC in every sensitivity analysis and the percentage improvement can increase up to 18%. We conclude that our initial calibration of the model does not provide the highest improvement. Our results suggest that increasing the time horizon or the level of transaction costs greatly increase the percentage improvement. The level of factor persistence strictly varies the total performance but the overall improvement seems to be similar across low and high mean reversion speeds. Finally, choosing a smaller value for relaxation probability leads to better performance, but we observe that values in the interval $0.1 \leq \nu \leq 0.3$ roughly provides the same objective values.

5. Dynamic Trading with Mean-Variance Objective

In this section, we analyze a dynamic trading problem for an investor with mean-variance preferences. Due to its foundational role in modern portfolio theory, mean-variance preferences have been widely studied in the multi-period portfolio choice literature. Basak and Chabakauri (2010) provides a survey of this literature and characterizes the optimal mean-variance portfolios under various stochastic investment opportunities. However, their economic setup does not include transaction costs or portfolio constraints. We provide this empirical experiment in order to illustrate that mean-variance preferences are accommodated in our general modeling framework and emphasize the potential large benefits of using optimal linear policy as opposed to LQC-based optimal trading rule.

Our model specification is inspired from the previous section and we follow the same model calibration described in Section 4.4. The main novelty in this application is the exact implementation of a mean-variance objective function. Although convex transaction costs and constraints can also be added to this objective function without losing any tractability in solving for the optimal linear policy, we will not do so here for the mere sake of comparing our approach with an approximate LQC policy in a simpler framework. Using the same calibration, we are interested in how to trade a single stock optimally over a short time horizon when its price changes can be predicted by two factors with different mean reversion speeds. We set our trading horizon to be 1 hour and implement trading decisions every 5 minutes which translates into a time horizon of $T = 12$ periods. In our simulation results, we also consider three other choices for time horizon, $T = 6$, $T = 18$, and $T = 24$, as robustness checks. Finally, we assume that the trader has zero shares as the initial position.

We have the same dynamics for price changes and factor realizations as in our previous

			Optimal	Projected		
			Linear	LQC	Difference	Improvement
			(\$k)	(\$k)	(\$k)	
Time Horizon	$T = 6$	Mean	-14.573	-14.577	0.0043	0.04%
		S.E.	0.105	0.105	0.0003	
	$T = 12$	Mean	6.13	5.73	0.40	6.98%
		S.E.	0.229	0.224	0.0095	
	$T = 18$	Mean	22.96	21.43	1.53	7.14%
		S.E.	0.339	0.349	0.028	
	$T = 24$	Mean	37.10	34.26	2.84	8.29%
		S.E.	0.443	0.459	0.059	
Transaction Costs	$\Lambda = 4.28 \times 10^{-6}$	Mean	24.01	22.73	1.28	5.63%
		S.E.	0.235	0.240	0.025	
	$\Lambda = 1.07 \times 10^{-5}$	Mean	15.84	14.74	1.10	7.46%
		S.E.	0.229	0.236	0.017	
	$\Lambda = 2.14 \times 10^{-5}$	Mean	6.13	5.73	0.40	6.98%
		S.E.	0.229	0.224	0.0095	
	$\Lambda = 2.67 \times 10^{-5}$	Mean	2.25	1.90	0.35	18.42%
		S.E.	0.222	0.225	0.0067	
Factor Persistence	$\Phi = \begin{bmatrix} 0.3573 & 0 \\ 0 & 0.0176 \end{bmatrix}$	Mean	10.51	9.94	0.57	8.21%
		S.E.	0.267	0.272	0.0127	
	$\Phi = \begin{bmatrix} 0.7146 & 0 \\ 0 & 0.0353 \end{bmatrix}$	Mean	6.13	5.73	0.40	6.98%
		S.E.	0.229	0.224	0.0095	
	$\Phi = \begin{bmatrix} 0.8000 & 0 \\ 0 & 0.0400 \end{bmatrix}$	Mean	5.52	5.06	0.46	9.09%
		S.E.	0.217	0.221	0.0082	
	$\Phi = \begin{bmatrix} 0.9000 & 0 \\ 0 & 0.0500 \end{bmatrix}$	Mean	4.32	3.93	0.39	9.92%
		S.E.	0.206	0.209	0.0072	
Relaxation Probability	$\nu = 0.1$	Mean	6.18	5.73	0.45	7.85%
		S.E.	0.233	0.224	0.0092	
	$\nu = 0.2$	Mean	6.13	5.73	0.40	6.98%
		S.E.	0.229	0.224	0.0095	
	$\nu = 0.3$	Mean	6.14	5.73	0.41	7.16%
		S.E.	0.233	0.228	0.0089	
	$\nu = 0.4$	Mean	5.89	5.73	0.16	2.79%
		S.E.	0.233	0.228	0.0088	

Table 3: Sensitivity analysis of results in the optimal execution example with respect to the main parameters of the model, i.e., length of time horizon, level of transaction costs, factor persistence and relaxation probability.

application:

$$f_{t+1} = (I - \Phi) f_t + \epsilon_{t+1}^{(1)}, \quad r_{t+1} = B f_t + \epsilon_{t+1}^{(2)},$$

where the noise terms are i.i.d., and normally distributed with zero-mean and covariance matrices given by $\text{Var}(\epsilon_{t+1}^{(1)}) = \Psi$ and $\text{Var}(\epsilon_{t+1}^{(2)}) = \Sigma$. We use the calibrated values in Section 4.4 for Φ , Σ , Ψ , and B . We set the coefficient of risk aversion γ to 5×10^{-4} . We use three other choices for Φ in our simulations in order to assess the robustness of our results with respect to mean reversion speeds.

We assume that the investor has mean-variance preferences and seeks to maximize expected terminal wealth subject to the penalty term from the variance of terminal wealth,

$$(25) \quad \underset{\pi \in \mathcal{U}_\pi}{\text{maximize}} \quad \mathbb{E}_\pi [W(\mathbf{x}, \mathbf{r})] - \frac{\gamma}{2} \text{Var}_\pi (W(\mathbf{x}, \mathbf{r})),$$

where $W(\mathbf{x}, \mathbf{r}) \triangleq W_0 + \sum_{t=1}^T x_t^\top r_{t+1}$ is the terminal wealth, and γ is the coefficient of risk aversion. The main difference in this objective function compared to the previous application of Section 4 is the replacement of per-period additive quadratic penalty terms in inventory with a single penalty term for the variance of the terminal wealth.

We find the optimal linear policy by solving the following optimization problem

$$(26) \quad \underset{(\mathbf{J}, \mathbf{d})}{\text{maximize}} \quad \sum_{t=1}^T \mathbb{E} \left[\left(d_t + \sum_{s=1}^t J_{s,t} f_s \right)^\top B f_t \right] - \frac{\gamma}{2} \text{Var} \left(\sum_{t=1}^T \left(d_t + \sum_{s=1}^t J_{s,t} f_s \right)^\top r_{t+1} \right),$$

where $x_t = d_t + \sum_{s=1}^t J_{s,t} f_s$ specifies the linear rebalancing rule. In Appendix A.3, we show that this program can be reduced to an exact deterministic convex optimization problem.

For comparison, we define an approximate LQC policy as follows. Since LQC cannot directly penalize for the variance of the terminal wealth, we consider an alternative LQC formulation that penalizes in a per period fashion according to

$$(27) \quad \underset{g}{\text{maximize}} \quad \mathbb{E} \left[\sum_{t=1}^T x_t^\top r_{t+1} - \frac{g}{2} x_t^\top \Sigma x_t \right].$$

Here, $g > 0$ is a risk aversion multiplier we will specify momentarily. It is easy to see that, in the absence of transaction costs, this problem is separable over time. Then, the optimal solution is myopic and is given by $x_t^* \triangleq (g\Sigma)^{-1} B f_t$. We then compute the best LQC policy by solving for the optimal value of the multiplier g that maximizes the true objective function

with the variance penalty,

$$(28) \quad \begin{aligned} & \underset{g}{\text{maximize}} && \mathbb{E} \left[\sum_{t=1}^T x_t^\top r_{t+1} \right] - \frac{\gamma}{2} \text{Var} \left(\sum_{t=1}^T x_t^\top r_{t+1} \right) \\ & \text{subject to} && x_t = (g\Sigma)^{-1} Bf_t, \quad 1 \leq t \leq T. \end{aligned}$$

Then, letting $x_t = (g^*\Sigma)^{-1} Bf_t$ provides an approximate LQC policy.

In Table 4 and Table 5, we report the simulation results from 50,000 trials illustrating the performance differential of optimal linear policy and LQC policy under four different choices of time horizon and mean reversion speeds for the return predicting factors¹⁰.

We observe that optimal linear policy outperforms the LQC policy substantially in every case and the percentage improvement in the objective value can increase up to 72%. Our results suggest that increasing the time horizon and the level of persistence in the signals is positively related with the corresponding percentage improvement.

6. Conclusion

This paper provides a highly tractable formulation for determining rebalancing rules in dynamic portfolio choice problems involving complex models of return predictability. Our rebalancing rule is a linear function of past return predicting factors and can be utilized in a wide spectrum of portfolio choice models with realistic considerations for risk measures, transaction costs, and trading constraints. We illustrate the broad utility of our method by showing its applicability across a broad range of modeling assumptions on these portfolio optimization primitives. As long as the underlying dynamic portfolio optimization problem is a convex programming problem (i.e., concave objective and convex decision constraints), the modified optimization problem seeking the optimal parameters of the linear decision rule will be a convex programming problem that is tractable numerically. We demonstrate in realistic numerical experiments that such modeling flexibility can offer significant practical benefits.

References

- R. F. Almgren and N. Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 3(2): 5–39, 2000.
- P. Balduzzi and A. W. Lynch. Transaction costs and predictability: some utility cost calculations. *Journal of Financial Economics*, 52(1):47–78, apr 1999.

¹⁰Our simulation results across different values of Σ and γ simply scale the objective value by the ratio of these parameter values. Hence, we omit a sensitivity analysis on these parameters.

		Optimal Linear	LQC	Increase
$T = 6$	Objective	703.8	632.6	11.3%
	S.E.	8.51	6.20	
	Average Wealth	1613.6	1235.5	30.6%
	S.E.	8.53	6.95	
	Variance of Wealth	3.64×10^6	2.41×10^6	50.9%
	S.E.	0.39×10^6	0.33×10^6	
$T = 12$	Sharpe Ratio	0.85	0.80	6.3%
	S.E.	0.005	0.005	
	Objective	1332.5	978.0	36.3%
	S.E.	13.45	9.92	
	Average Wealth	2913.7	1919.5	60.6%
	S.E.	11.25	8.68	
$T = 18$	Variance of Wealth	6.32×10^6	3.36×10^6	88.0%
	S.E.	0.58×10^6	0.48×10^6	
	Sharpe Ratio	1.16	0.99	17.1%
	S.E.	0.006	0.006	
	Objective	1916.7	1234	55.3%
	S.E.	17.60	10.75	
$T = 24$	Average Wealth	4196.4	2300.3	82.4%
	S.E.	13.51	9.24	
	Variance of Wealth	9.12×10^6	4.27×10^6	113.8%
	S.E.	0.72×10^6	0.59×10^6	
	Sharpe Ratio	1.39	1.11	24.8%
	S.E.	0.006	0.005	
$T = 24$	Objective	2532.3	1472.8	71.9%
	S.E.	22.7	14.2	
	Average Wealth	5532.6	2814.8	96.6%
	S.E.	15.49	10.36	
	Variance of Wealth	12.0×10^6	5.37×10^6	123.6%
	S.E.	0.94×10^6	0.76×10^6	
	Sharpe Ratio	1.60	1.21	31.5%
	S.E.	0.006	0.006	

Table 4: Summary of the performance statistics in the mean-variance preferences example of the optimal linear policy and the LQC policy under different choices of time horizon. For each policy, we report the resulting objective value, average wealth, variance of wealth and Sharpe ratio estimates along with their corresponding standard errors.

		Optimal Linear	LQC	Increase
$\Phi = \begin{bmatrix} 0.3573 & 0 \\ 0 & 0.0176 \end{bmatrix}$	Objective	1316.3	931.6	41.3%
	S.E.	13.71	8.46	
	Average Wealth	2945.5	1727.5	70.5%
	S.E.	11.42	7.98	
	Variance of Wealth	6.52×10^6	3.18×10^6	104.7%
	S.E.	0.62×10^6	0.47×10^6	
$\Phi = \begin{bmatrix} 0.7146 & 0 \\ 0 & 0.0353 \end{bmatrix}$	Objective	1332.5	978.0	36.3%
	S.E.	13.45	9.92	
	Average Wealth	2913.7	1919.5	60.6%
	S.E.	11.25	8.68	
	Variance of Wealth	6.32×10^6	3.36×10^6	88.0%
	S.E.	0.58×10^6	0.48×10^6	
$\Phi = \begin{bmatrix} 0.8000 & 0 \\ 0 & 0.0400 \end{bmatrix}$	Objective	1336.8	987.8	35.3%
	S.E.	13.45	8.96	
	Average Wealth	2916.4	1842.7	58.3%
	S.E.	11.24	8.27	
	Variance of Wealth	6.32×10^6	3.42×10^6	84.8%
	S.E.	0.57×10^6	0.48×10^6	
$\Phi = \begin{bmatrix} 0.9000 & 0 \\ 0 & 0.0500 \end{bmatrix}$	Objective	1345.7	1024.7	31.3%
	S.E.	13.49	9.24	
	Average Wealth	2931.8	1914.1	53.2%
	S.E.	11.26	8.43	
	Variance of Wealth	6.34×10^6	3.56×10^6	78.3%
	S.E.	0.57×10^6	0.49×10^6	
	Sharpe Ratio	1.16	1.01	14.7%
	S.E.	0.006	0.006	

Table 5: Summary of the performance statistics in the mean-variance preferences example of the optimal linear policy and the LQC policy under different choices of factor persistence. For each policy, we report the resulting objective value, average wealth, variance of wealth and Sharpe ratio estimates along with their corresponding standard errors.

- S. Basak and G. Chabakauri. Dynamic mean-variance asset allocation. *Review of Financial Studies*, 23(8):2970–3016, 2010. doi: 10.1093/rfs/hhq028.
- A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99(2):351–376, 2004.
- A. Ben-Tal, B. Golany, A. Nemirovski, and J. Vial. Retailer-supplier flexible commitments contracts: A robust optimization approach. *Manufacturing & Service Operations Management*, 7: 248–271, July 2005.
- D. P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, 2nd edition, 2000. ISBN 1886529094.
- D. Bertsimas and V. Goyal. On the power and limitations of affine policies in two-stage adaptive optimization. *Mathematical Programming*, pages 1–41, 2011.
- D. Bertsimas and A. Lo. Optimal control of execution costs. *Journal of Financial Markets*, 1(1): 1–50, 1998. ISSN 1386-4181.
- D. Bertsimas, D. A. Iancu, and P. A. Parrilo. Optimality of affine policies in multistage robust optimization. *Math. Oper. Res.*, 35, May 2010.
- S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, March 2004. ISBN 0521833787.
- S. Boyd, M. T. Mueller, B. O’Donoghue, and Wang Y. Performance bounds and suboptimal policies for multi-period investment. Working paper, 2012.
- M. W. Brandt and P. Santa-Clara. Dynamic portfolio selection by augmenting the asset space. *The Journal of Finance*, 61(5):2187–2217, 2006.
- M. W. Brandt, A. Goyal, P. Santa-Clara, and J. R. Stroud. A simulation approach to dynamic portfolio choice with an application to learning about return predictability. *Review of Financial Studies*, 18(3):831–873, 2005.
- M. W. Brandt, P. Santa-Clara, and R. Valkanov. Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *Review of Financial Studies*, 22(9):3411–3447, September 2009.
- D. B. Brown and J. E. Smith. Dynamic portfolio optimization with transaction costs: Heuristics and dual bounds. Working paper, 2010.
- G. C. Calafiore. An affine control method for optimal dynamic asset allocation with transaction costs. *SIAM J. Control Optim.*, 48:2254–2274, June 2009.
- J. Y. Campbell and L. M. Viceira. *Strategic asset allocation*. Oxford University Press Oxford, 2002.
- N. Chen, R. Roll, and S. A. Ross. Economic forces and the stock market. *The Journal of Business*, 59(3):pp. 383–403, 1986.
- X. Chen, M. Sim, and P. Sun. A robust optimization perspective on stochastic programming. *Operations Research*, 55:1058–1071, November 2007.

- X. Chen, M. Sim, P. Sun, and J. Zhang. A linear decision-based approximation approach to stochastic programming. *Operations Research*, 56:344–357, March 2008.
- G. M. Constantinides. Capital market equilibrium with transaction costs. *The Journal of Political Economy*, 94(4):pp. 842–862, August 1986.
- J. Cvitanic. Theory of portfolio optimization in markets with frictions. In E. Jouini, J. Cvitanic, and Marek Musiela, editors, *Handbooks in Mathematical Finance*. Cambridge University Press, 2001.
- M. H. A. Davis and A. R. Norman. Portfolio selection with transaction costs. *Mathematics of Operations Research*, 15(4):676–713, 1990.
- V. V. Desai, V. F. Farias, and C. C. Moallemi. Pathwise optimization for linear convex systems. *Working paper*, 2011.
- J. B. Detemple, R. Garcia, and M. Rindisbacher. A Monte Carlo method for optimal portfolios. *The Journal of Finance*, 58(1):401–446, February 2003.
- B. Dumas and E. Luciano. An exact solution to a dynamic portfolio choice problem under transactions costs. *The Journal of Finance*, 46(2):pp. 577–595, 1991.
- E. F. Fama and K. R. French. Multifactor explanations of asset pricing anomalies. *The Journal of Finance*, 51(1):pp. 55–84, 1996.
- N. Gârleanu and L. H. Pedersen. Dynamic trading with predictable returns and transaction costs. *The Journal of Finance*, 68(6):2309–2340, 2013.
- J. Gatheral. No-dynamic-arbitrage and market impact. *Quantitative Finance*, 10(7):749–759, 2010.
- P. Glasserman and X. Xu. Robust portfolio control with stochastic factor dynamics. Working paper, 2011.
- W. N. Goetzmann and P. Jorion. Testing the predictive power of dividend yields. *Journal of Finance*, 48(2):663–679, 1993.
- M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 1.21. <http://cvxr.com/cvx>, April 2011.
- R. C. Grinold and R. N. Kahn. *Active Portfolio Management*. McGraw-Hill, 2nd edition, 1999.
- S. J. Grossman and Z. Zhou. Optimal investment strategies for controlling drawdowns. *Mathematical Finance*, 3(3):241–276, 1993.
- S. L. Heston, R. A. Korajczyk, and R. Sadka. Intraday patterns in the cross-section of stock returns. *Journal of Finance*, 65(4):1369–1407, 2010.
- M. Hora. Tactical liquidity trading and intraday volume. Working Paper, 2006.
- D. Kuhn, W. Wiesemann, and A. Georghiou. Primal and dual linear decision rules in stochastic and robust optimization. *Mathematical Programming*, pages 1–33, 2009.

- H. Liu. Optimal consumption and investment with transaction costs and multiple risky assets. *Journal of Finance*, 59(1):289–338, 02 2004.
- H. Liu and M. Loewenstein. Optimal portfolio selection with transaction costs and finite horizons. *Review of Financial Studies*, 15(3):805–835, 2002.
- A. W. Lynch and P. Balduzzi. Predictability and transaction costs: The impact on rebalancing rules and behavior. *Journal of Finance*, 55(5):2285–2309, October 2000.
- A. W. Lynch and S. Tan. Multiple risky assets, transaction costs, and return predictability: Allocation rules and implications for U.S. investors. *Journal of Financial and Quantitative Analysis*, 45(04):1015–1053, August 2010.
- Robert C. Merton. Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory*, 3(4):373–413, December 1971.
- A. Nemirovski, A. Juditsky, G. Lan, and A. Shapiro. Stochastic approximation approach to stochastic programming. *SIAM Journal on Optimization*, 19(4):1574–1609, 2009.
- A. A. Obizhaeva and J. Wang. Optimal trading strategy and supply/demand dynamics. Working paper, 2005.
- A. Shapiro. Monte Carlo sampling methods. In A. Ruszczyński and A. Shapiro, editors, *Stochastic Programming*, volume 10 of *Handbooks in Operations Research and Management Science*, pages 353 – 425. Elsevier, 2003.
- A. Shapiro and A. Nemirovski. On complexity of stochastic programming problems. In Panos M. Pardalos, Donald Hearn, Vaithilingam Jeyakumar, and Alexander Rubinov, editors, *Continuous Optimization*, volume 99 of *Applied Optimization*, pages 111–146. Springer, 2005.
- S. E. Shreve and H. M. Soner. Optimal investment and consumption with transaction costs. *The Annals of Applied Probability*, 4(3):pp. 609–692, 1994.
- J. Skaf and S. P. Boyd. Nonlinear Q-design for convex stochastic control. *IEEE Transactions on Automatic Control*, 54(10):2426–2430, October 2009.
- J. Skaf and S. P. Boyd. Design of affine controllers via convex optimization. *IEEE Transactions on Automatic Control*, 55(11):2476–2487, November 2010.

Technical Appendices

A. Efficient Exact Formulations

In this section, we will provide efficient exact formulations of dynamic portfolio choice problems using the class of linear policies for our feasible set of policies. In particular, we will consider a number of the examples of dynamic portfolio choice problems discussed in Section 2.1. These examples include features such as constraints on portfolio holdings, transaction costs, and risk measures. In each case, we will demonstrate how the optimization problem (11) can be transformed into a deterministic convex program by explicit analytical evaluation of the objective function $\mathbf{E}[p(\cdot, \mathbf{f})]$ and the constraint set \mathcal{C} .

Exact formulations require the evaluation of expectations taken over the sample path of factor realizations \mathbf{f} . In order to do this, we will make the following assumption for the rest of this section:

Assumption 4 (Gaussian factors). *Assume that the sample path \mathbf{f} of factor realizations is jointly Gaussian. In particular, denote by $F_t \triangleq (f_1, \dots, f_t)^\top \in \mathbb{R}^{Kt}$ the vector of all factors observed by time t . We assume that $F_t \sim N(\theta_t, \Omega_t)$, where $\theta_t \in \mathbb{R}^{Kt}$ is the mean vector and $\Omega_t \in \mathbb{R}^{Kt \times Kt}$ is the covariance matrix.*

Note that this assumption is not necessary for the practical application of our method. Indeed, as discussed in Section 3, sample average approximation or stochastic approximation methods can be used to determine the optimal policy in the absence of such an assumption. However, under Assumption 4, the optimal linear policy can be solved via a deterministic convex optimization program that does not involve any sampling. This is often more computationally efficient as well as more accurate (since there are no errors introduced by sampling).

With this assumption, the trades of any linear policy will also be jointly normally dis-

tributed, as each such policy is affine transformations of the factors. Formally, let

$$(29) \quad M_t \triangleq \begin{bmatrix} E_{1,t} & E_{2,t} & \dots & E_{t,t} \end{bmatrix} \in \mathbb{R}^{N \times Kt}$$

be the matrix of time t policy coefficients, so that the trade vector is given by $u_t = c_t + M_t F_t$. With this representation, it is easy to see that $u_t \sim N(\bar{u}_t, V_t)$, where the mean vector and covariance matrix are given by

$$(30) \quad \bar{u}_t \triangleq \mathbf{E}[u_t] = c_t + M_t \theta_t, \quad V_t \triangleq \text{Var}(u_t) = M_t \Omega_t M_t^\top.$$

Similarly, the portfolio x_t at time t is normally distributed. we have that

$$(31) \quad x_t = x_0 + \sum_{i=1}^t u_i = x_0 + \sum_{i=1}^t \left(c_i + \sum_{s=1}^i E_{s,i} f_s \right) = d_t + \sum_{s=1}^t J_{s,t} f_s,$$

where $d_t \triangleq x_0 + \sum_{i=1}^t c_i$ and $J_{s,t} \triangleq \sum_{\ell=s}^t E_{s,\ell}$. With this representation, it is easy see that $x_t \sim N(\kappa_t, Y_t)$, where

$$(32) \quad \kappa_t \triangleq \mathbf{E}[x_t] = d_t + P_t \theta_t, \quad Y_t \triangleq \text{Var}(x_t) = P_t \Omega_t P_t^\top,$$

$$(33) \quad P_t \triangleq \begin{bmatrix} J_{1,t} & J_{2,t} & \dots & J_{t,t} \end{bmatrix}.$$

A.1. Linear Constraints

We will provide formulations for linear equality or inequality constraints on trades or positions, in the context of linear rebalancing policies. These types of constraints appear frequently in portfolio choice due to regulatory reasons such as short sale restriction, liquidation purposes or diversification needs (e.g., keeping a specific industry exposure under a certain limit).

A.1.1. Equality Constraints

Equality constraints appear often in portfolio choice, particularly in portfolio execution problems when the investor needs to liquidate a certain portfolio (i.e., $x_T = \mathbf{0}$) or constructs a certain target portfolio by the end of the time horizon (i.e., $x_T = \bar{x}$).

Suppose that for some time t , we have a linear equality constraint on the trade vector u_t , of the form $Au_t = b$. Here, $A \in \mathbb{R}^{M \times N}$ and $b \in \mathbb{R}^N$. This constraint can be written as

$$(34) \quad Ac_t + AM_tF_t = b.$$

Under Assumption 4, the left side of (34) is normally distributed. Therefore, for (34) to hold almost surely, we must have that the left side has mean b and zero covariance. Thus, we require that

$$(35) \quad Ac_t = b, \quad AM_t = \mathbf{0}.$$

Thus, the linear equality constraint (34) on the trade vector u_t is equivalent to the linear equality constraint (35) on the policy coefficients (c_t, M_t) . Linear equality constraints on the portfolio position x_t can be handled similarly.

A.1.2. Inequality Constraints

Inequality constraints on trades or positions are common as well. One example is a short-sale constraint, which would require that $x_t \geq \mathbf{0}$ for all times t . When the factor realizations do not have bounded support, inequality constraints cannot be enforced almost surely. This is true in the Gaussian case: there is a chance, however small, that factors may take extreme values, and if the policy is a linear function of the factors, this may cause an inequality constraint to be violated. In other words, under Assumption 4 with linear constraints, any non-deterministic linear rebalancing policy will typically be infeasible. Note that while this

is immediate in the exact formulation, this difficulty remains if we employ other solution methods such as sample average approximation. In the SAA approach, for example, any non-deterministic linear rebalance policy will become infeasible given enough samples.

In order to account for such constraints in a linear rebalancing policy, instead of enforcing inequality constraints almost surely, we will enforce them at a given level of confidence. For example, given a vector $a \in \mathbb{R}^N$ and a scalar b , instead of enforcing the linear constraint $a^\top u_t \leq b$, almost surely, we can consider a relaxation where we seek to guarantee that it is violated with small probability. In other words, we can impose the chance constraint $\mathbb{P}(a^\top u_t > b) \leq \eta$, for a small value of the parameter η . The following lemma illustrates that this can be accomplished explicitly:

Lemma 1. *Given $\eta \in [0, 1/2]$, a non-zero vector $a \in \mathbb{R}^N$, and a scalar b , the chance constraint $\mathbb{P}(a^\top u_t > b) \leq \eta$ is equivalent to the constraint*

$$a^\top (c_t + M_t \theta_t) - b + \Phi^{-1}(1 - \eta) \left\| \Omega_t^{1/2} M_t^\top a \right\|_2 \leq 0$$

on the policy coefficients (c_t, M_t) , where $\Phi^{-1}(\cdot)$ is the inverse cumulative normal distribution.

Proof. This proof follows standard arguments in convex optimization (see, e.g., Boyd and Vandenberghe, 2004). Let \bar{u}_t and V_t be the mean and the variance of u_t as given in (30). Then,

$$\mathbb{P}(a^\top u_t > b) = \mathbb{P}(\beta_t + \sigma_t Z > 0),$$

where

$$\beta_t \triangleq a^\top \bar{u}_t - b, \quad \sigma_t \triangleq \|V_t^{1/2} a\|_2 \neq 0,$$

and Z is a standard normal random variable. Thus,

$$\mathbb{P}(a^\top u_t > b) = 1 - \Phi(-\beta_t/\sigma_t).$$

Note that the this probability is less than or equal to η if and only if

$$\beta_t + \Phi^{-1}(1 - \eta)\sigma_t \leq 0.$$

Substituting (30) into definitions for β_t and σ_t , we obtain the desired result. ■

A similar approach can be applied to incorporate linear inequality constraints on the portfolio position x_t with high confidence.

In many situations (e.g., short-sale constraints), it may not be sufficient to enforce an inequality constraint only probabilistically. In such cases, when a linear rebalancing policy is applied, the resulting trades can be projected onto the constraint set so as to ensure that the constraints are always satisfied. When the linear policy is designed, however, it is helpful to incorporate the desired constraints probabilistically so as to account for their presence. We will demonstrate this idea in the application in Section 4.

A.2. Transaction Costs

In this section, we will provide efficient exact formulations for the transaction cost functions discussed in Section 2.1, in the context of linear rebalancing policies. In general, one might consider a total transaction cost of

$$\text{TC}(\mathbf{u}) \triangleq \sum_{t=1}^T \text{TC}_t(u_t)$$

for executing the sample path of trades \mathbf{u} , where $\text{TC}_t(u_t)$ is the cost of executing the trade vector u_t at time t . As seen in Section 2.1, we typically wish to subtract an expected transaction cost term from investor's objective. Hence efficient exact formulations for transaction costs involve explicit analytical computation of $\mathbf{E}[\text{TC}(\mathbf{u})] = \sum_{t=1}^T \mathbf{E}[\text{TC}_t(u_t)]$, when each trade vector u_t is specified by a linear policy.

Under a linear policy, $u_t \sim N(\bar{u}_t, V_t)$ is distributed as a normal random variable with

mean and covariance, (\bar{u}_t, V_t) , specified with the policy coefficients, (\mathbf{E}, \mathbf{c}) , through (30). Then, the evaluation of expected transaction costs reduces to the evaluation of the expected value of the per-period transaction cost function $\text{TC}_t(\cdot)$ for a Gaussian random variable. This can be handled on a case-by-case basis as follows:

- **Quadratic transaction costs.** In the case of quadratic transaction costs, as seen in Example 2, the per period transaction cost function is given by $\text{TC}_t(u_t) \triangleq \frac{1}{2} u_t^\top \Lambda u_t$, where $\Lambda \in \mathbb{R}^{N \times N}$ is a positive definite matrix. In this case, $\mathbb{E}[\text{TC}_t(u_t)] = \frac{1}{2} (\bar{u}_t \Lambda \bar{u}_t + \text{tr}(\Lambda V_t))$.
- **Proportional transaction costs.** In the case of proportional transaction costs, as discussed in Example 4, the per period transaction cost function is given by

$$\text{TC}_t(u_t) \triangleq \sum_{i=1}^N \chi_i |u_{t,i}|,$$

where $\chi_i > 0$ is a proportionality constant specific to security i . Using the properties of the folded normal distribution, we obtain

$$\mathbb{E}[\text{TC}_t(u_t)] = \sum_{i=1}^N \chi_i \left(\sqrt{\frac{2V_{t,i}}{\pi}} \exp \left\{ -\frac{\bar{u}_{t,i}^2}{2V_{t,i}} \right\} + \bar{u}_{t,i} \left\{ 1 - 2\Phi \left(-\frac{\bar{u}_{t,i}}{\sqrt{V_{t,i}}} \right) \right\} \right),$$

where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal random variable.

- **Nonlinear transaction costs.** In the case of nonlinear transaction costs, as discussed in Example 4, the per period transaction cost function is given by

$$\text{TC}_t(u_t) \triangleq \sum_{i=1}^N \chi_i |u_{t,i}|^\beta,$$

where $\chi_i > 0$ is a proportionality constant specific to security i , and $\beta \geq 1$ is an exponent capturing the degree of nonlinearity. As in the proportional case, evaluating

the Gaussian expectation explicitly results in

$$\mathbb{E}[\text{TC}_t(u_t)] = \sum_{i=1}^N \chi_i \Gamma\left(\frac{1+\beta}{2}\right) \frac{(2V_{t,i})^{\frac{\beta}{2}}}{\sqrt{\pi}} {}_1F_1\left(-\frac{\beta}{2}; \frac{1}{2}; -\frac{\bar{u}_{t,i}^2}{2V_{t,i}}\right),$$

where $\Gamma(\cdot)$ is the gamma function and ${}_1F_1(\cdot)$ is the confluent hypergeometric function of the first kind.

A.3. Terminal Wealth and Risk Aversion

In many of the portfolio choice examples in Section 2.1, an investor wishes to maximize expected wealth net of transaction costs, subject to a penalty for risk, i.e.,

$$(36) \quad \underset{\pi \in \mathcal{U}_{\mathbb{F}}}{\text{maximize}} \mathbb{E}_{\pi} \left[W(\mathbf{x}, \mathbf{r}) - \text{TC}(\mathbf{u}) - \text{RA}(\mathbf{x}, \mathbf{f}, \mathbf{r}) \right].$$

Here, $W(\cdot)$ is the terminal wealth associated with a sample path, $\text{TC}(\cdot)$ are the transaction costs, and $\text{RA}(\cdot)$ is a penalty for risk aversions. Exact calculation of expected transaction costs for linear policies were discussed in Section A.2. Here, we will discuss exact calculation of the expected terminal wealth and the risk aversion penalty.

To begin, note that the terminal wealth depends on realized returns in addition to factor realizations. Hence, we will make the following assumption:

Assumption 5 (Gaussian returns). *As in Example 2, assume that for each time $t \geq 0$, returns evolve according to*

$$(37) \quad r_{t+1} = \mu_t + Bf_t + \epsilon_{t+1}^{(2)},$$

where μ_t is a deterministic vector, B is a matrix of factor loadings, and $\epsilon_t^{(2)}$ are zero-mean i.i.d. Gaussian disturbances with covariance Σ .

Note that the critical assumption we are making here is that the factor realizations \mathbf{f}

and the sample path of security returns \mathbf{r} are jointly Gaussian. The particular form (37) is chosen out of convenience but is not necessary.

We can calculate the expected terminal wealth as

$$\begin{aligned} \mathbb{E}[W(\mathbf{x}, \mathbf{r})] &= W_0 + \sum_{t=1}^T \mathbb{E}[x_t^\top r_{t+1}] = W_0 + \sum_{t=1}^T \left(\mu_t^\top \kappa_t + \mathbb{E}[x_t^\top B f_t] \right), \\ &= W_0 + \sum_{t=1}^T \left(\mu_t^\top \kappa_t + d_t^\top B \delta_t + \sum_{s=1}^t \left(\delta_s^\top (B(I - \Phi)^{t-s} J_{s,t}^\top) \delta_s + \text{tr} \left(B(I - \Phi)^{t-s} J_{s,t}^\top \omega_s \right) \right) \right), \end{aligned}$$

where ω_s is the s th $K \times K$ diagonal block matrix of Ω_t .

For the risk aversion penalty, we consider two cases:

- **Per period risk penalty.** Consider risk aversion penalties that decompose over time as

$$\text{RA}(\mathbf{x}, \mathbf{f}, \mathbf{r}) = \sum_{t=1}^N \text{RA}_t(x_t),$$

where $\text{RA}_t(\cdot)$ is a function which penalizes for risk aversion based on the positions held at time t . One such case is the quadratic penalty $\text{RA}_t(x_t) \triangleq \frac{\gamma}{2} x_t^\top \Sigma x_t$ of Example 2, where $\gamma > 0$ is a risk penalty proportionality constant. Here, the investor seeks to penalize in proportion to the conditional per period variance of the portfolio value. So long as the expectation of $\text{RA}_t(\cdot)$ can be calculated for Gaussian arguments, then the overall expected risk aversion penalty can be calculated exactly. This can be accomplished for a variety of functions. For example, quadratic penalties can be handled in a manner analogous to the quadratic transaction costs discussed in Section A.2.

- **Terminal wealth risk penalty.** Alternatively, as discussed in Example 5, a more natural risk aversion criteria might be to penalize risk as a function of the *terminal* wealth. Specifically, an investor with mean-variance preferences would consider a risk aversion penalty $\text{RA}(\mathbf{x}, \mathbf{f}, \mathbf{r}) \triangleq -\frac{\gamma}{2} \text{Var}(W(\mathbf{x}, \mathbf{r}))$, where $\gamma > 0$ is a risk penalty proportionality constant.

Following the notation of Section A, we will compute $\text{Var}[W(\mathbf{x}, \mathbf{r})]$ analytically and

demonstrate that the resulting expression is a quadratic convex function of the policy coefficients. Without loss of generality, assume that $W_0 = 0$ and $\mu_t = \mathbf{0}$. Let $d \in \mathbb{R}^{NT \times 1}$, $e \in \mathbb{R}^{NT \times 1}$, $L \in \mathbb{R}^{NT \times KT}$, and $D \in \mathbb{R}^{NT \times KT}$ with

$$d \triangleq \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_{T-1} \\ d_T \end{bmatrix}, \quad e \triangleq \begin{bmatrix} \epsilon_1^{(2)} \\ \epsilon_2^{(2)} \\ \vdots \\ \epsilon_{T-1}^{(2)} \\ \epsilon_T^{(2)} \end{bmatrix}, \quad L \triangleq \begin{bmatrix} B & 0 & \dots & 0 & 0 \\ 0 & B & & \dots & 0 \\ \vdots & & \ddots & \dots & 0 \\ & \dots & & B & 0 \\ 0 & \dots & & 0 & B \end{bmatrix},$$

$$D \triangleq \begin{bmatrix} J_{1,1} & 0 & \dots & 0 & 0 \\ J_{1,2} & J_{2,2} & & \dots & 0 \\ \vdots & & \ddots & \dots & 0 \\ & \dots & & \dots & 0 \\ J_{1,T} & \dots & & J_{T-1,T} & J_{T,T} \end{bmatrix}.$$

Then, observe that $\begin{bmatrix} x_1^\top & \dots & x_T^\top \end{bmatrix}^\top = d + DF_T$, and $W(\mathbf{x}, \mathbf{r}) = (d + DF_T)^\top (LF_T + e)$.

Using the independence between e and F_T and $\mathbb{E}[e] = 0$, we obtain

$$\begin{aligned} \text{Var}(W(\mathbf{x}, \mathbf{r})) &= \mathbb{E} \left[(d + DF_T)^\top (LF_T + e)(LF_T + e)^\top (d + DF_T) \right] \\ &\quad - \mathbb{E} \left[(d + DF_T)^\top (LF_T + e) \right]^2 \\ (38) \quad &= \mathbb{E} \left[(d + DF_T)^\top ee^\top (d + DF_T) \right] \\ &\quad + \mathbb{E} \left[(d + DF_T)^\top (LF_T)(LF_T)^\top (d + DF_T) \right] \\ &\quad - \mathbb{E} \left[(d + DF_T)^\top (LF_T) \right]^2 \end{aligned}$$

We need the following fact from multivariate statistics in order to compute the expectations.

Fact 1. *If z is a random vector with mean m and variance S , and Q, A, H are constant*

matrices, and a is a constant vector, then

$$\begin{aligned}\mathbb{E} \left[z^\top Q z \right] &= \text{tr}(QS) + m^\top Q m, \\ \mathbb{E} \left[(Az + a)^\top H z (Hz)^\top (Az + a) \right] &= \text{tr} \left(AS(H^\top A + A^\top H)SH^\top \right) \\ &\quad + \left((Am + a)^\top H + (Hm)^\top A \right) S \left((Am + a)^\top H + (Hm)^\top A \right)^\top \\ &\quad + \left(\text{tr} \left(ASH^\top \right) + (Am + a)^\top H m \right)^2.\end{aligned}$$

We will consider each of the three terms in (38) separately. The first expectation, $\mathbb{E} \left[(d + DF_T)^\top ee^\top (d + DF_T) \right]$, can be evaluated using the independence of $\epsilon_s^{(2)}$ and $\epsilon_t^{(2)}$ when $s \neq t$ and Fact 1.

$$\begin{aligned}\mathbb{E} \left[(d + DF_T)^\top ee^\top (d + DF_T) \right] &= \sum_{t=1}^T \mathbb{E} \left[x_t^\top \mathbb{E} \left[\epsilon_{t+1}^{(2)} \left(\epsilon_{t+1}^{(2)} \right)^\top \mid \mathcal{F}_t \right] x_t \right] \\ &= \sum_{t=1}^T \mathbb{E} \left[x_t^\top \Sigma x_t \right] \\ &= \sum_{t=1}^T (d_t + P_t \theta_t)^\top \Sigma (d_t + P_t \theta_t) + \text{tr} \left(\Sigma P_t \Omega_t P_t^\top \right).\end{aligned}$$

Note that this term is equivalent to the sum of per-period risk penalties considered in the previous section.

The remaining expectations in (38) can be directly computed using Fact 1 with replacing z with F_T where $F_T \sim N(\Theta_T, \Omega_T)$.

$$\begin{aligned}\mathbb{E} \left[(d + DF_T)^\top (LF_T)(LF_T)^\top (d + DF_T) \right] &= \text{tr} \left(D\Omega_T(L^\top D + D^\top L)\Omega_T L^\top \right) \\ &\quad + \left((D\Theta_T + d)^\top L + (L\Theta_T)^\top D \right) \Omega_T \left((D\Theta_T + d)^\top L + (L\Theta_T)^\top D \right)^\top \\ &\quad + \left(\text{tr} \left(D\Omega_T L^\top \right) + (D\Theta_T + d)^\top L \Theta_T \right)^2.\end{aligned}$$

Finally,

$$\mathbf{E} \left[(d + DF_T)^\top (LF_T) \right]^2 = \left(\text{tr} \left(D\Omega_T L^\top \right) + (D\Theta_T + d)^\top L\Theta_T \right)^2.$$

Summing all three terms in (38), we obtain the exact expression for the variance of terminal wealth:

$$(39) \quad \begin{aligned} \text{Var} (W(\mathbf{x}, \mathbf{r})) &= \sum_{t=1}^T (d_t + P_t \theta_t)^\top \Sigma (d_t + P_t \theta_t) + \text{tr} \left(\Sigma P_t \Omega_t P_t^\top \right) \\ &\quad + \text{tr} \left(D\Omega_T (L^\top D + D^\top L) \Omega_T L^\top \right) \\ &\quad + \left((D\Theta_T + d)^\top L + (L\Theta_T)^\top D \right) \Omega_T \left((D\Theta_T + d)^\top L + (L\Theta_T)^\top D \right)^\top. \end{aligned}$$

Note that this expression is convex in our decision variables, as expected.

B. Derivation of the LQC Policies

We can derive a closed-form solution for our trading policy when the problem satisfies the LQC framework. We guess a functional form for the value function and show that this functional form is preserved at each time step.

Using dynamic programming principle and $u_t = (x_t - x_{t-1})$, the value function $V_t(x_{t-1}, f_t)$ satisfies

$$V_{t-1}(x_{t-1}, f_t) = \underset{x_t}{\text{maximize}} \left(x_t^\top (Bf_t) - \frac{1}{2} (x_t - x_{t-1})^\top \Lambda (x_t - x_{t-1}) + \mathbf{E}[V_t(x_t, f_{t+1})] \right).$$

We guess the following quadratic form for our value function:

$$V_t(x_t, f_{t+1}) = -\frac{1}{2} x_t^\top A_{xx,t} x_t + x_t^\top A_{xf,t} f_{t+1} + \frac{1}{2} f_{t+1}^\top A_{ff,t} f_{t+1} + \frac{1}{2} m_t.$$

Then,

$$\mathbb{E}[V_t(x_t, f_{t+1})] = -\frac{1}{2}x_t^\top A_{xx,t}x_t + x_t^\top A_{xf,t}(I - \Phi) f_t + \frac{1}{2}f_t^\top (I - \Phi)^\top A_{ff,t}(I - \Phi) f_t + \frac{1}{2}(\text{tr}(\Psi A_{ff,t}) + m_t).$$

At the last period, we need $x_T = 0$, and our value function equals

$$V_{T-1}(x_{T-1}, f_t) = -\frac{1}{2}x_{T-1}^\top \Lambda x_{T-1}$$

which satisfies our functional form with

$$A_{xx,T-1} = \Lambda \quad A_{xf,T-1} = \mathbf{0} \quad A_{ff,T-1} = \mathbf{0} \quad m_{T-1} = 0.$$

For all $t < T - 1$, we maximize the quadratic objective $-\frac{1}{2}x_t^\top Q_t x_t + x_t^\top q_t + b_t$ where

$$Q_t = \Lambda + A_{xx,t}$$

$$q_t = \Lambda x_{t-1} + (B + A_{xf,t}(I - \Phi)) f_t$$

$$b_t = -\frac{1}{2}x_{t-1}^\top \Lambda x_{t-1} + \frac{1}{2}f_t^\top (I - \Phi)^\top A_{ff,t}(I - \Phi) f_t + \text{tr}(\Psi A_{ff,t}) + m_t$$

Then, the optimal x_t is given by $Q_t^{-1}q_t$ and x_t and u_t are given by

$$x_t = (\Lambda + A_{xx,t})^{-1} (\Lambda x_{t-1} + (B + A_{xf,t}(I - \Phi)) f_t)$$

$$u_t = (\Lambda + A_{xx,t})^{-1} (\Lambda x_{t-1} + (B + A_{xf,t}(I - \Phi)) f_t) - x_{t-1}$$

The maximum then occurs at $\frac{1}{2}q_t^\top Q_t^{-1}q_t + b_t$ and we obtain the following recursions:

$$\begin{aligned}
A_{xx,t-1} &= -\Lambda (\Lambda + A_{xx,t})^{-1} \Lambda + \Lambda \\
A_{xf,t-1} &= \Lambda (\Lambda + A_{xx,t})^{-1} (B + A_{xf,t} (I - \Phi)) \\
A_{ff,t-1} &= (B + A_{xf,t} (I - \Phi))^\top (\Lambda + A_{xx,t})^{-1} (B + A_{xf,t} (I - \Phi)) + (I - \Phi)^\top A_{ff,t} (I - \Phi) \\
m_{t-1} &= \text{tr}(\Psi A_{ff,t}) + m_t
\end{aligned}$$

Using these recursions, we can compute the optimal expected payoff of the dynamic program. Using $f_0 = N(0, \Omega_0)$,

$$\begin{aligned}
\mathbb{E}[V_0(x_0, f_1)] &= \mathbb{E}[\mathbb{E}[V_0(x_0, f_1)|f_0]] \\
&= \mathbb{E}\left[-\frac{1}{2}x_0^\top A_{xx,0}x_0 + x_0^\top A_{xf,0} (I - \Phi) f_0 + \frac{1}{2}f_0^\top (I - \Phi)^\top A_{ff,0} (I - \Phi) f_0 + \frac{1}{2}(\text{tr}(\Omega_0 A_{ff,0}) + m_0)\right] \\
&= -\frac{1}{2}x_0^\top A_{xx,0}x_0 + \frac{1}{2}\left(\text{tr}\left(\Omega_0 (I - \Phi)^\top A_{ff,0} (I - \Phi)\right) + \sum_{t=0}^{T-2} \text{tr}(\Psi A_{ff,t})\right).
\end{aligned}$$

C. Exact Formulation of Best Linear Execution Policy

We will first compute the expectation in the objective of (21) and write the equivalent deterministic form. We will then replace probabilistic constraints with deterministic constraints using Lemma 1, and finally obtain the deterministic version of the stochastic program in (21).

We start working with the expectation in the objective function. For each t , we have to compute the expectation of the following two terms, $\mathbb{E}\left[x_t^\top (B f_t)\right]$, and $\mathbb{E}\left[u_t^\top \Lambda u_t\right]$. First, we derive the statistics for f_t , u_t and x_t . We first note that

$$f_t = (I - \Phi)^T e_0 + \sum_{s=1}^t (I - \Phi)^{t-s} \epsilon_s^{(1)}.$$

Letting $F_t \triangleq (f_1, \dots, f_t)^\top$, Then, in vector form, we have the following representation

$$F_t = \begin{bmatrix} (I - \Phi)f_0 \\ (I - \Phi)^2 f_0 \\ \vdots \\ (I - \Phi)^{t-1} f_0 \\ (I - \Phi)^t f_0 \end{bmatrix} + \underbrace{\begin{bmatrix} I & 0 & \dots & 0 & 0 \\ (I - \Phi) & I & 0 & \dots & 0 \\ \vdots & (I - \Phi) & \ddots & \dots & 0 \\ (I - \Phi)^{t-1} & \dots & \ddots & I & 0 \\ (I - \Phi)^t & \dots & & (I - \Phi) & I \end{bmatrix}}_{\triangleq A_t} \begin{bmatrix} \epsilon_1^{(1)} \\ \epsilon_2^{(1)} \\ \vdots \\ \epsilon_{t-1}^{(1)} \\ \epsilon_t^{(1)} \end{bmatrix}.$$

Using this representation, we compute the mean

$$(40) \quad \theta_t \triangleq \mathbb{E}[F_t] = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_{t-1} \\ \delta_t \end{bmatrix} \triangleq \begin{bmatrix} (I - \Phi)f_0 \\ (I - \Phi)^2 f_0 \\ \vdots \\ (I - \Phi)^{t-1} f_0 \\ (I - \Phi)^t f_0 \end{bmatrix},$$

and the covariance matrix

$$(41) \quad \Omega_t \triangleq \text{Var}[F_t] = A_t \begin{bmatrix} \Psi & 0 & \dots & 0 & 0 \\ 0 & \Psi & & \dots & 0 \\ \vdots & & \ddots & \dots & 0 \\ & \dots & & \Psi & 0 \\ 0 & \dots & & 0 & \Psi \end{bmatrix} A_t^\top.$$

Note that Ω_t is a block diagonal matrix with t blocks of size $K \times K$. Recall that in Section A, we defined

$$(42) \quad M_t \triangleq \begin{bmatrix} E_{1,t} & E_{2,t} & \dots & E_{t,t} \end{bmatrix}$$

Then, $u_t = c_t + M_t F_t$ and we have the following moments for u_t :

$$(43) \quad \begin{aligned} \mu_t &\triangleq \mathbb{E}[u_t] = c_t + M_t \theta_t \\ V_t &\triangleq \text{Var}(u_t) = M_t \Omega_t M_t^\top. \end{aligned}$$

Therefore, u_t is normally distributed with mean μ_t and covariance matrix V_t . Similarly, we can obtain the statistics for x_t . Using (32),

$$\begin{aligned} \kappa_t &\triangleq \mathbb{E}[x_t] = d_t + P_t \theta_t \\ Y_t &\triangleq \text{Var}(x_t) = P_t \Omega_t P_t^\top. \end{aligned}$$

Using Fact 1, we can compute each term in the expectation.

$$\begin{aligned} \mathbb{E} \left[x_t^\top (B f_t) \right] &= \mathbb{E} \left[d_t^\top B f_t + \sum_{s=1}^t f_s^\top J_{s,t}^\top B f_t \right] \\ &= d_t^\top B \delta_t + \sum_{s=1}^t \mathbb{E} \left[f_s^\top J_{s,t}^\top B \mathbb{E}[f_t | f_s] \right] \\ &= d_t^\top B \delta_t + \sum_{s=1}^t \mathbb{E} \left[f_s^\top J_{s,t}^\top B (I - \Phi)^{t-s} f_s \right] \\ &= d_t^\top B \delta_t + \sum_{s=1}^t \left(\delta_s^\top (B (I - \Phi)^{t-s} J_{s,t}^\top) \delta_s + \text{tr} \left(B (I - \Phi)^{t-s} J_{s,t}^\top \omega_s \right) \right) \end{aligned}$$

where ω_s is the s th diagonal block matrix of Ω_t having a size of $K \times K$. Finally, for the transaction cost terms,

$$\begin{aligned} \mathbb{E} \left[u_t^\top \Lambda u_t \right] &= \mathbb{E} \left[(c_t + M_t F_t)^\top \Lambda (c_t + M_t F_t) \right] \\ &= (c_t + M_t \theta_t)^\top \Lambda (c_t + M_t \theta_t) + \text{tr} \left(\Lambda M_t \Omega_t M_t^\top \right) \end{aligned}$$

Summing up all the terms, our final objective function in deterministic form equals

$$\begin{aligned} \underset{c_t, E_{s,t}}{\text{maximize}} \sum_{t=1}^T \left\{ d_t^\top B \delta_t + \sum_{s=1}^t \left(\delta_s^\top (B(I - \Phi)^{t-s} J_{s,t}^\top) \delta_s + \text{tr} \left(B(I - \Phi)^{t-s} E_{s,t}^\top \omega_s \right) \right) \right. \\ \left. + \frac{1}{2} \left((c_t + M_t \theta_t)^\top \Lambda (c_t + M_t \theta_t) + \text{tr} \left(\Lambda M_t \Omega_t M_t^\top \right) \right) \right\}, \end{aligned}$$

which is a quadratic function of the policy parameters.

We now rewrite the equality constraint, $x_T = 0$ in terms of policy parameters. In order to enforce this equality, we need

$$d_T = 0 \text{ and } J_{s,T} = \mathbf{0} \quad s = 1, \dots, T.$$

Lastly, we replace probabilistic constraints with deterministic constraints using Lemma 1. Note that $\mathbf{P}(x_t \leq 0) \leq \eta$ can be written as $\mathbf{P}(-x_t \geq 0) \leq \eta$. Then, using Lemma 1,

$$(-d_t - P_t \theta_t) + \Phi^{-1}(1 - \eta) \left\| \left(P_t \Omega_t P_t^\top \right)^{1/2} \right\|_2 \leq 0.$$

Similarly, we obtain that $\mathbf{P}(u_t \geq 0) \leq \eta$ can be replaced by

$$(c_t + M_t \theta_t) + \Phi^{-1}(1 - \eta) \left\| \left(M_t \Omega_t M_t^\top \right)^{1/2} \right\|_2 \leq 0.$$

Combining all the results, we obtain the deterministic version of the stochastic program in (21), a second-order cone program:

$$\begin{aligned}
(44) \quad & \underset{c_t, E_{s,t}}{\text{maximize}} \sum_{t=1}^T \left\{ d_t^\top B \delta_t + \sum_{s=1}^t \left(\delta_s^\top (B(I - \Phi)^{t-s} J_{s,t}^\top) \delta_s + \text{tr} \left(B(I - \Phi)^{t-s} E_{s,t}^\top \omega_s \right) \right) \right. \\
& \quad \left. + \frac{1}{2} \left((c_t + M_t \theta_t)^\top \Lambda (c_t + M_t \theta_t) + \text{tr} \left(M_t^\top \Lambda M_t \Omega_t \right) \right) \right\} \\
& \text{subject to } d_t = x_0 + \sum_{i=1}^t c_i \quad t = 1, \dots, T, \\
& \quad J_{s,t} = \sum_{i=s}^t E_{s,i} \quad 1 \leq s \leq t \leq T, \\
& \quad (-d_t - P_t \theta_t) + \Phi^{-1}(1 - \eta) \left\| \left(P_t \Omega_t P_t^\top \right)^{1/2} \right\|_2 \leq 0 \quad t = 1, \dots, T, \\
& \quad (c_t + M_t \theta_t) + \Phi^{-1}(1 - \eta) \left\| \left(M_t \Omega_t M_t^\top \right)^{1/2} \right\|_2 \leq 0 \quad t = 1, \dots, T, \\
& \quad d_T = 0 \text{ and } J_{s,T} = \mathbf{0}.
\end{aligned}$$

Note that the number of decision variables is greater than that of the original execution problem in (16). Total number of decision variables in a problem with N securities, K factors and T periods equals $2NT + NKT(T + 1)$ which is on the order of $O(NKT^2)$.